

**PROGRAMA DE PESQUISA
ESCOLA DE CONTAS E GESTÃO – ECG-TCE/RJ**

RELATÓRIO FINAL

Título do Projeto:

CONTROLE EXTERNO BASEADO EM ANÁLISE DE DADOS: ESTÃO OS TRIBUNAIS DE
CONTAS PREPARADOS?

Grupo de Pesquisa:

Pesquisador-Coordenador: BRUNO MATTOS SOUZA DE SOUZA MELO

Pesquisador-Avaliador: ALBERTO DE FONTES TAVARES NETO

Processo administrativo: TCE-RJ Nº 302.782-6/15

ESCOLA DE CONTAS E GESTÃO – ECG-TCE/RJ

Diretor-Geral

JOÃO PAULO MENEZES LOURENÇO

Coordenadora-Geral de Estudos e Pesquisas

ROSA MARIA CHAISE

Comissão Permanente de Estudos e Pesquisas – COPEP

ROSA MARIA CHAISE (Presidente)

ALBERTO DE FONTES TAVARES NETO

ÁLVARO GUILHERME MIRANDA

CARLOS LEANDRO DOS SANTOS REGINALDO

CLÁUDIO MARTINELLI MURTA

HEITOR DELGADO CORREA

JEAN MARCEL DE FARIA NOVO

JOÃO PAULO MENEZES LOURENÇO

JORGE HENRIQUE MUNIZ DA CONCEIÇÃO

MARIANA MIRANDA AUTRAN SAMPAIO

NINA QUINTANILHA ARAUJO

PAULO CÉSAR BESSA NEVES

SÉRGIO PAULO VIEIRA VILAÇA

SÉRGIO LINO DA SILVA CARVALHO

Pesquisador-avaliador

Data

28/08/2017

SUMÁRIO

1 - Introdução

- 1.1 Justificativa e relevância
- 1.2 Questão norteadora
- 1.3 Objetivos
- 1.4 Referencial teórico

2 - Metodologia

- 2.1 Considerações preliminares
- 2.2 O que é uma escala de Likert?
- 2.3 Construção de uma escala de Likert
- 2.4 Vantagens e desvantagens da escala de Likert
- 2.5 Avaliação dos resultados da escala
- 2.6 Etapas da pesquisa
 - 2.6.1 Definição conceitual do construto a ser investigado
 - 2.6.2 Geração dos itens e validação com especialistas
 - 2.6.3 População, amostra e seleção dos entrevistados
 - 2.6.4 Pré-teste em amostra piloto
 - 2.6.5 Aplicação da escala nos tribunais

3 - Resultados / Conclusões

4 - Referências

5 - Apêndices

- Apêndice 1 - Cronograma previsto e executado
- Apêndice 2 - Dificuldades encontradas
- Apêndice 3 - – Formulário para validação dos itens por especialistas
- Apêndice 4 - Quantidade de Tribunais participantes da pesquisa e quantidade de respostas obtidas

6 - Parecer do pesquisador-avaliador

1 - INTRODUÇÃO

1.1 Justificativa e relevância

É amplamente reconhecido que o uso de tecnologia da informação tem o condão de tornar a atuação das instituições mais eficiente. E não poderia ser diferente com os Tribunais de Contas. Também estas instituições têm a obrigação de ser eficientes. É o que determina a Constituição Federal ao submeter a administração pública ao princípio da eficiência.

Com o aumento da disponibilidade de dados e o barateamento dos recursos computacionais, o uso de *analytics* surge como uma forma efetiva de as corporações adquirirem diferencial competitivo (Davenport e Harris, 2007). De acordo com relatório da empresa SAS denominado *How to Compete on Analytics: The Analytical Center of Excellence*, (SAS, 2008) vivemos na era do *analytics* e esta é a nova arma secreta das empresas.

Situação semelhante ocorre com organizações governamentais, que devem entregar à sociedade melhores serviços a custos menores. No âmbito dos Tribunais de Contas isso se traduz, entre outras coisas, em maior eficiência na realização do controle.

Reconhecendo a importância do uso de análise de dados como instrumento indutor de eficiência, os Tribunais de Contas (TCs) implantaram a Rede Nacional de Informações Estratégicas para o Controle Externo - Infocontas, que tem por objetivo, entre outras coisas, fazer uso de dados para direcionar a escolha das fiscalizações a serem realizadas, priorizando as que apresentem um maior risco (Infocontas, 2013).

1.2 Problema de pesquisa

Havendo a necessidade de se tornar a atuação dos Tribunais de Contas mais eficiente, é desejável que eles façam o melhor uso possível dos dados disponíveis para a tomada de decisões. Isto será possível caso o TC possua um ambiente favorável à adoção de uma estratégia de controle que faça uso intensivo de análise de dados.

Isto posto, a questão que surge é: **Estão os Tribunais de Contas preparados para a implantação de uma estratégia de controle baseada em análise de dados?**

1.3 Objetivos

O objetivo geral da pesquisa é mensurar a maturidade analítica dos tribunais de contas brasileiros e sua adequação à adoção de uma estratégia de controle baseada em análise de dados.

Como objetivos específicos, a pesquisa se propôs a:

- Criar uma escala de maturidade analítica aplicável a órgãos de controle;
- Mensurar o nível de maturidade analítica dos tribunais com base na escala criada;
- Definir os contornos e requisitos de uma estratégia de controle baseada em análise de dados;
- Avaliar a adequação do nível de maturidade analítica observada nos TCs com uma estratégia de controle fundamentada em análise de dados;
- Identificar oportunidade de melhorias no nível de maturidade analítica dos TCs

pesquisados e, notadamente, no TCE-RJ.

1.4 Referencial Teórico

Davenport e Harris (2007) definem Inteligência Analítica da seguinte forma:

Inteligência Analítica é a utilização extensiva de dados, análises quantitativas e estatísticas, modelos explicativos e preditivos e gestão baseada em fatos para orientar decisões e ações.

A definição aponta para aspectos essenciais que devem ser objeto de investigação junto aos Tribunais de Contas com vistas à mensuração de sua **maturidade/capacidade analítica**, entendida como o nível de inteligência analítica existente na instituição.

O primeiro aspecto diz respeito à **utilização extensiva de dados**. Este é um aspecto fundamental. O uso extensivo de dados traz questões relativas à existência, qualidade e fidedignidade dos dados. Trata-se de pré-requisito essencial, sem o qual o TC não pode avançar na maturidade analítica.

Outro aspecto trata do uso de **análises quantitativas e estatísticas**. De nada servirão os dados se não existirem técnicos capazes de realizar sua análise. Isso implica o conhecimento tanto de técnicas quantitativas como de *softwares* analíticos.

A utilização de **modelos estatísticos explicativos e preditivos** assume papel relevante quando se procura por relações de causalidade entre variáveis (modelos explicativos) ou se deseja realizar previsões (modelos preditivos) a partir de dados já disponíveis no Tribunal.

De nada adianta realizar análise de dados e possuir bons modelos preditivos se as decisões continuarem a ser tomadas com base unicamente na experiência do gestor, e não se considerar os fatos concretos. Daí surge a importância de se possuir uma **gestão baseada em fatos**, que se caracteriza pela tomada de decisões e implementação de ações tomando por base os dados disponíveis.

O Tribunal de Contas da União (TCU 2015) estabelece entre as prioridades da gestão, para o biênio 2015-2016, o **Controle Baseado em Análise de Dados** e elenca as iniciativas estratégicas que contribuirão para a realização deste objetivo:

- Aprimorar ferramenta (*dashboard*) para utilização em fiscalização contínua na área de educação;
- Utilizar ferramenta de georreferenciamento para avaliação de risco de empreendimentos de infraestrutura;
- Elaborar metodologias de classificação de risco mediante análise preditiva;
- Fortalecer a auditoria financeira no TCU;
- Aperfeiçoar e utilizar metodologia de análise de risco para subsidiar a seleção de objetos de controle;
- Capacitar o corpo técnico nos conceitos trazidos pela nova Contabilidade Pública, bem como nas ferramentas gerenciais (Tesouro Gerencial) destinadas ao tratamento de dados

de execução financeiro-orçamentária;

- Tratar dados de bases informacionais a fim de orientar ações de controle;
- Conceber soluções de TI para análise automática do conteúdo de informações não estruturadas relacionadas aos processos de controle externo.

São oito iniciativas que abrangem um amplo espectro de ações que passam por compreensão dos dados disponíveis (auditoria financeira, conceitos de contabilidade pública), tomada de decisão com base em fatos (metodologia de análise de risco, tratar dados de bases informacionais) uso de *software* (*dashboards*, ferramenta de georrefenciamento, solução para análise automática de informações não estruturadas) e de técnicas quantitativas (análise preditiva).

A capacitação dos técnicos é um elemento essencial e perpassa todas as iniciativas.

O fato de estas ações constarem do planejamento estratégico da Instituição já sinaliza uma predisposição da organização com sua consecução.

2 - METODOLOGIA

2.1 Considerações preliminares

A mensuração do conceito de capacidade ou maturidade analítica terá por base a elaboração de uma escala de Likert, cujos itens foram definidos com fundamento no referencial teórico descrito na Introdução, os quais foram posteriormente validados junto a especialistas.

Antes de se apresentar as etapas metodológicas da pesquisa, entende-se necessário apresentar uma justificativa para a escolha da escala de Likert para mensurar maturidade analítica.

O primeiro aspecto a ser observado é que a maturidade analítica é um conceito, e dessa forma, sua mensuração de forma direta é impossível. Não obstante, mediante a definição de variáveis diretamente observáveis associadas ao conceito é possível se obter uma forma de se mensurar este conceito. Numa escala de Likert, as variáveis são os itens elaborados para compor a escala, que uma vez criada é aplicada à população-alvo da pesquisa, no presente caso, os técnicos das instituições a serem avaliadas (Tribunais de Contas).

Vieira (2011) apresenta um grande conjunto de escalas de Likert que se destinam a avaliar uma ampla gama de fenômenos de interesse na área de *marketing*: comportamento do consumidor, qualidade, lealdade, satisfação, envolvimento, emoções, marca e escalas para desempenho empresarial.

A esta altura, uma questão já deverá estar na mente dos leitores: se a escala vai aferir a percepção dos técnicos que trabalham no Tribunal de Contas, como medir a maturidade analítica dos TCs propriamente ditos?

O raciocínio aqui é que, se um determinado Tribunal de Contas realmente possui um elevado nível de maturidade analítica e, se os itens estão de fato aferindo os aspectos relevantes do construto, os técnicos que lá trabalham terão uma percepção positiva deste fato e, assim, se considerarmos que o escore de um Tribunal seja a média dos escores dos técnicos que participaram da pesquisa, será possível avaliar a maturidade do Tribunal em função da percepção de seus técnicos.

Destaca-se que essa abordagem se diferencia do que usualmente tem sido a prática em mensurar maturidade. A abordagem predominante consiste em identificar as condições que propiciam alto nível de maturidade, avaliar os aspectos operacionais que se relacionam com estas condições e definir estágios de maturidade de acordo com os aspectos operacionais desejados ou observados em cada estágio.

Usualmente, as condições operacionais existentes são avaliadas por intermédio de entrevistas ou via questionário encaminhado a pessoas-chaves.

Acredita-se que a abordagem psicométrica adotada tem a vantagem de aferir o quanto os aspectos operacionais efetivamente estão sendo percebidos pelos seus destinatários. Não são

raros os casos em que, apesar de uma instituição possuir determinados recursos, estes são desconhecidos ou subutilizados ou existem apenas formalmente.

2.2 O que é uma escala de Likert?

Posto de uma forma simples, uma escala tipo Likert consiste de uma coleção de itens Likert.

Um item likert consiste de uma assertiva e uma escala de resposta, isto é, um conjunto de opções de respostas oferecido aos respondentes. Um exemplo de um item likert seria:

Assertiva (Item)	Opções de Resposta
As iniciativas de análise de dados são isoladas.	<input type="checkbox"/> Discordo completamente
	<input type="checkbox"/> Discordo
	<input type="checkbox"/> Não concordo nem discordo
	<input type="checkbox"/> Concordo
	<input type="checkbox"/> Concordo completamente

De acordo com Cunha (2007), “Uma escala tipo Likert é composta por um conjunto de frases (itens) em relação a cada uma das quais se pede ao sujeito que está a ser avaliado para manifestar o grau de concordância desde o discordo totalmente (nível 1), até o concordo totalmente (nível 5, 7 ou 11)”.

Esta escala foi desenvolvido por Rensis Likert com o objetivo de mensurar atitudes no contexto das ciências comportamentais. Consiste em tomar um construto e desenvolver um conjunto de afirmações relacionadas à sua definição, para as quais os respondentes emitirão seu grau de concordância.

A escala para mensurar a maturidade analítica nos Tribunais de Contas (construto avaliado) será um subconjunto dos itens que inicialmente serão gerados após as análises de confiabilidade que deverão ser realizadas.

2.3 Construção de uma escala de Likert

De acordo com Moser e Kalton (1971), as etapas envolvidas na elaboração de uma escala são as seguintes: (a) reunir um conjunto inicial de itens, a partir do qual aqueles que figurarão na escala serão selecionados; (b) escolher, dentre os itens do conjunto inicial, aqueles que farão parte da escala final.

A escolha dos itens que permanecerão é feita a partir de um estudo exploratório, do qual a análise fatorial exploratória, apresentada mais adiante, é um componente relevante.

Os itens devem ser elaborados de forma a diferenciar aqueles que tenham uma atitude favorável ao objeto de estudo, daqueles que tenham uma atitude desfavorável, razão pela qual os itens para os quais praticamente todos respondam da mesma forma são insatisfatórios. Também é importante que os itens busquem avaliar todos os aspectos da atitude em estudo. No caso da pesquisa, os itens buscam avaliar 3 aspectos que estão relacionados à maturidade analítica: o componente organizacional, que tem a ver com características atribuíveis à Instituição, o componente humano, que tem a ver com as competências técnicas existentes, e o componente

tecnológico, que tem a ver com a disponibilidade e o efetivo uso de dados, *software* e *hardware*.

2.4 Vantagens e desvantagens da escala de Likert

Dentre as vantagens de se utilizar a escala de Likert, tem-se a facilidade de construção do questionário, possibilitando ao entrevistado que expresse a intensidade de sua opinião em perguntas complexas, além de ser bastante conveniente para a construção de questionários *on-line*, permitindo ao entrevistado fazer comparações entre os itens, modificações e ajustes à sua resposta com facilidade.

Por outro lado, dentre os inconvenientes da escala há o fato de que duas pessoas que tenham expressado opiniões diferentes em cada um dos itens da escala podem obter o mesmo valor final. Há também uma dificuldade em trabalhar com respostas neutras, como "não concordo nem discordo". Outro ponto negativo é o fato de os entrevistados tenderem a concordar com as declarações apresentadas, fato conhecido como viés de concordância.

2.5 Avaliação dos resultados da escala

Estando de posse de uma escala que tenha sido aplicada a um conjunto de pessoas, como devem ser avaliados os resultados?

Uma vez que a cada opção em um item de Likert é atribuído um escore (1 – Discordo completamente, 2- Discordo, 3-Não discordo nem concordo, 4 – Concordo, 5 – Concordo completamente) a atitude de cada pessoa é obtida pela soma dos escores que ela atribuiu a cada item. Em razão desta característica, esta escala também é conhecida por escala de soma.

Tendo-se decidido pelo uso da escala de Likert, descreve-se a seguir as etapas da pesquisa.

2.6 Etapas da pesquisa

Tomando por base as orientações de Vieira (2011), a pesquisa seguiu as etapas a seguir elencadas:

- Definição do conceito a ser investigado. No caso da presente pesquisa, o conceito de maturidade analítica dos Tribunais de Contas;
- Geração dos itens, com base na literatura e/ou pesquisa de campo;
- Validação de conteúdo com especialistas;
- Aplicação de pré-teste e realização de ajustes no instrumento de coleta (análise fatorial exploratória, confiabilidade)
- Aplicação do questionário final e análise dos dados (análise fatorial exploratória e confirmatória, confiabilidade, dimensionalidade, definição da quantidade de fatores)

2.6.1 Definição conceitual do construto a ser investigado

O objetivo principal da pesquisa é criar uma escala para a mensuração da maturidade analítica dos Tribunais de Contas. Assim, algo que se deve fazer é definir o objeto da mensuração.

Na introdução foi definida maturidade analítica ou capacidade analítica (nesta pesquisa consideram-se ambos os termos indistintamente) como sendo o nível de inteligência analítica

existente na instituição. Adota-se também o conceito de inteligência analítica dado por Davenport e Harris (2007) e apresentado na introdução.

2.6.2 Geração dos itens e validação com especialistas

A partir do referencial teórico pesquisado, descrito na introdução, em especial Davenport e Harris (2007), foi gerado um conjunto de itens iniciais para compor o questionário, reunindo os três fatores com base nos quais os citados autores propõem a avaliação da maturidade analítica de uma instituição. Estes fatores são os seguintes:

- fatores organizacionais;
- fatores humanos; e
- fatores tecnológicos.

Foram elaborados 22 itens que foram submetidos à validação de especialistas, mediante um formulário que lhes foi encaminhado por *e-mail*. Cópia deste formulário consta do Apêndice 3.

Forma encaminhados formulários a técnicos de instituições de controle com experiência no uso de ferramentas analíticas para a realização de auditorias.

Os especialistas que responderam à solicitação concordaram com os fatores utilizados e fizeram sugestões que ajudaram na melhoria dos itens.

Os itens gerados e a indicação dos fatores que buscam mensurar foram os seguintes:

Quadro 01 – Relação dos itens

Código	Descrição	Fator
Q01	O Tribunal dispõe de técnicos capacitados em análise de dados para a realização de auditorias/fiscalizações.	Humano
Q02	Nos dois últimos anos o Tribunal realizou alguma auditoria/fiscalização em que utilizou técnicas de análise de dados.	Humano
Q03	O Tribunal oferece a seus técnicos capacitação em análise de dados.	Organizacional
Q04	As iniciativas de uso de análise de dados são isoladas.	Organizacional
Q05	Inexiste incentivo por parte do Tribunal para que os técnicos se capacitem em análise de dados.	Organizacional
Q06	O Tribunal utiliza dashboard para monitorar métricas chaves para o controle externo.	Tecnológico
Q07	Além dos dados obtidos com os jurisdicionados, o Tribunal também disponibiliza aos técnicos informações geradas por outras instituições.	Tecnológico
Q08	O Tribunal dispõe de um setor que tem como missão gerenciar informações para dar suporte às atividades de controle externo do Tribunal.	Organizacional
Q09	A análise de risco é feita de forma subjetiva, valendo-se da experiência do técnico.	Organizacional
Q10	A análise de dados é utilizada para auxiliar na detecção de fraudes.	Organizacional
Q11	O Tribunal possui em sua estrutura organizacional um setor encarregado de garantir a qualidade dos dados utilizados para a realização das auditorias/fiscalizações.	Organizacional

Q12	Inexiste apoio da alta administração do Tribunal ao uso de análise de dados para realizar a atividade de controle.	Organizacional
Q13	Os recursos computacionais (hardware) disponibilizados aos técnicos são inadequados para trabalhar com grandes bases de dados.	Tecnológico
Q14	Recursos de geoprocessamento ainda não são utilizados na realização de auditorias/fiscalizações.	Tecnológico
Q15	Os técnicos não se sentem confortáveis para trabalhar com grandes bases de dados.	Humano
Q16	A seleção de objetos de fiscalização é realizada com base na experiência dos técnicos, sem o auxílio de análise de risco.	Organizacional
Q17	O processo de tomada de decisões no Tribunal é baseado em evidência.	Organizacional
Q18	O Tribunal possui ações estratégicas previstas explicitamente em seu Planejamento Estratégico visando aumentar sua capacidade analítica.	Organizacional
Q19	O uso de softwares que suportam análise de dados (SAS, R, Stata, Python, WEKA, KNIME, SQL etc.) é difundido entre os auditores.	Humano
Q20	Os técnicos têm dificuldade de dispor dos dados necessários à realização de auditorias/fiscalização.	Tecnológico
Q21	O uso de mineração de texto para auxiliar no exercício do controle externo ainda não é uma realidade no Tribunal.	Organizacional
Q22	O Tribunal dispõe de técnicos capacitados em softwares analíticos nos setores do controle externo que realizam auditorias/fiscalizações.	Humano

Considerando que os itens elaborados comporão a escala de Likert a ser utilizada, buscou-se seguir as orientações dadas por Rob Jonhs (2010):

- (a) evitar itens que mencionem mais de uma atitude;
- (b) evitar itens que tenham referência quantitativa;
- (c) evitar questões dúbias; e
- (d) evitar o viés de concordância.

Moser e Kalton (1971) recomendam também evitar itens excessivamente complexos, itens ambíguos, itens envolvendo dupla negativa e itens vagos.

O viés de concordância é a tendência de se concordar com as afirmações em alguma medida, independentemente de seu conteúdo, o que constitui uma fragilidade intrínseca à construção de itens tipo Likert. (Rob Jonhs, 2010).

Objetivando mitigar a ocorrência do viés de concordância, realizou-se o balanceamento do conjunto de itens. Especificamente, das 22 questões que constaram do questionário inicial, 10 foram balanceadas. (Questões Q04, Q05, Q09, Q12, Q13, Q14, Q15, Q16, Q20, Q21). O balanceamento consiste em formular a afirmativa de forma que a resposta mais favorável à atitude avaliada esteja em sentido oposto à de uma questão não balanceada.

Optou-se por uma escala de 5 pontos em razão de pesquisas que têm confirmado que dados

oriundos de escalas de Likert tornam-se significativamente menos acuradas quando o número de pontos na escala é inferior a 5 ou superior a 7. Esses estudos, por outro lado, não fornecem uma base para que se prefira uma escala de 5 pontos a uma de 7 pontos (Rob John, 2010)

O questionário foi aplicado utilizando a ferramenta LimeSurvey (<https://www.limesurvey.org/>) disponibilizada pela Diretoria Geral de Informática.

O agrupamento dos itens por fator é apresentada a seguir:

Quadro 2 – Agrupamento dos itens por fatores

Fator	Questões
Organizacional	
Processos	Q08,Q09,Q10,Q16,Q21
Patrocínio	Q04,Q12
Capacitação	Q03,Q05
Estrutura	Q11,Q17
Cultura	Q18
Humano	
Habilidades	Q01,Q02,Q15,Q19,Q22
Tecnológico	
Software	Q06,Q14
Dados	Q07,Q20
Hardware	Q13

2.6.3 População, amostra e seleção dos entrevistados

População objeto da pesquisa consiste dos técnicos que atuam em auditoria e fiscalização nos 34 Tribunais de Contas existentes no país.

A amostra-piloto consistiu nos técnicos que atuam em auditoria/fiscalização no Tribunal de Contas do Estado do Rio de Janeiro.

Para a amostra final, como não dispúnhamos de uma relação com os técnicos de cada Tribunal objeto da pesquisa, que atuam na área de auditoria/fiscalização, a abordagem utilizada para se obter os respondentes foi encaminhar o questionário a pelo menos um técnico em cada Tribunal e solicitar que o questionário fosse repassado ao máximo de pessoas possível. Esta abordagem é conhecida por amostragem *snowball*.

2.6.4 Pré-teste em amostra-piloto

Após a etapa de validação dos itens junto a especialistas, a etapa seguinte consistiu em aplicar um questionário a uma amostra-piloto contendo os 22 itens iniciais (acima elencados) com vistas à sua validação. A amostra-piloto consistiu em encaminhar, por *e-mail*, o questionário a 120 técnicos do Tribunal de Contas do Estado do Rio de Janeiro. Os técnicos selecionados foram os que participaram de uma capacitação em auditoria governamental oferecida pela Secretaria-Geral de Controle Externo – SGE.

A análise dos dados coletados foi realizada utilizando o *software* estatístico R e os pacotes `psych`, `likert`, `ggplot2`, `tydir`, `corrplot`, `readxl`. Esta análise consta do **Apêndice 5**.

A análise consistiu em calcular as correlações entre itens, realizar uma análise fatorial exploratória com vistas a identificar eventuais itens que deveriam ser excluídos e realizar a verificação da consistência interna utilizando o alfa de Cronbach.

Foi conduzida uma análise de componentes principais (PCA) em 49 observações utilizando-se tanto a rotação varimax quanto oblimin. Esta última, realizada em razão de se considerar existir relação entre os fatores, apresentou melhor configuração para a interpretação dos fatores.

O índice de Kaiser-Meyer-Olkin (KMO), utilizado para indicar a adequação da amostra à realização da análise, foi de 0.51, estando no limiar do aceitável de acordo com Hair et al. (2009). O teste de esfericidade de Bartlett, (qui-quadrado 383,36 e p-valor < 1,150913e-09) indicou que as correlações entre os itens são suficientemente altas para a realização da PCA com vistas à extração dos fatores.

Uma análise inicial foi realizada para se obter os autovalores para as componentes existentes nos dados. Sete componentes apresentaram autovalores superiores ao critério de Kaiser e, em conjunto, explicam 66% da variância. O *screeplot* também mostrou a mesma quantidade de fatores a ser extraída.

Examinando-se os resultados obtidos, conclui-se que:

O **Fator 1** recebeu contribuições significativas dos seguintes itens:

Q06 - O Tribunal utiliza dashboard para monitorar métricas chaves para o controle

Q04 - As iniciativas de uso de análise de dados são isoladas

Q21 - O uso de mineração de texto para auxiliar no exercício do controle externo ainda não é uma realidade

Q19 - O uso de softwares que suportam análise de dados é difundido entre os auditores

Q22 - O Tribunal dispõe de técnicos capacitados em softwares analíticos

Este fator pode ser interpretado como se referindo à **existência de pessoal capacitado em softwares e técnicas analíticas**.

O **Fator 2** recebeu contribuições significativas dos seguintes itens:

Q08 - O Tribunal possui um setor que tem como missão gerenciar informações para dar suporte

às atividades de controle externo do Tribunal

Q11 - O Tribunal possui em sua estrutura organizacional um setor encarregado de garantir a qualidade dos dados utilizados para a realização das auditorias/fiscalizações

Q10 - A análise de dados é utilizada para auxiliar na detecção de fraudes

Q07 - Além dos dados obtidos com os jurisdicionados, o Tribunal também disponibiliza aos técnicos dados obtidos de outras instituições

Este fator parece estar associado à **existência e uso de dados**.

O **Fator 3** recebeu contribuições significativas dos seguintes itens:

Q05 - Inexiste incentivo por parte do Tribunal para que os técnicos se capacitem em análise de dados

Q12 - Inexiste apoio da alta administração do Tribunal ao uso de análise de dados para realizar a atividade de controle

Q03 - O Tribunal oferece a seus técnicos capacitação em análise de dados

Este fator se relaciona a **apoio institucional para o uso de análise de dados**.

O **Fator 4** recebeu contribuições significativas do seguinte item:

Q13 - Os recursos computacionais (hardware) disponibilizados aos técnicos são inadequados para trabalhar com grandes bases de dados

Este fator se relaciona à **infraestrutura inadequada para manipular grandes bases de dados**

O **Fator 5** recebeu contribuição significativa dos seguintes itens:

Q18 - O Tribunal possui ações estratégicas previstas explicitamente em seu Planejamento Estratégico visando aumentar sua capacidade analítica

Q20 - Os técnicos têm dificuldade de dispor dos dados necessários à realização de auditorias/fiscalizações

Q09 - A análise de risco é feita de forma subjetiva, valendo-se da experiência do técnico

Este fator parece não ter uma interpretação muito clara, mas talvez seja possível vinculá-lo à **vontade institucional para o uso de análise de dados**.

O **Fator 6** recebeu contribuição significativa do seguinte item:

Q15 - Os técnicos não se sentem confortáveis para trabalhar com grandes bases de dados

Este fator pode ser vinculado à **competência técnica para manuseio de grandes bases de dados**.

O **Fator 7** recebeu contribuição significativa do seguinte item:

Q16 - A seleção de objetos de fiscalização é realizada com base na experiência dos técnicos, sem o auxílio de análise de risco

Este fator poderia indicar a **tomada de decisões baseada em fatos**.

Com isso, seriam removidas da escala final as seguintes questões: Q01, Q02, Q14, Q17.

Em resumo tem-se a seguinte composição dos fatores:

Fator	Itens
1	Q04, Q06, Q19, Q21, Q22
2	Q07, Q08, Q10, Q11
3	Q03, Q05, Q12
4	Q13
5	Q09, Q18, Q20
6	Q15
7	Q16

2.6.5 Aplicação da escala nos tribunais

Construída a escala, o passo seguinte é sua aplicação aos técnicos dos demais Tribunais de Contas com vistas a derivar a pontuação de cada um deles na escala, que refletirá o seu nível de maturidade analítica, conforme percebido pelos técnicos que lá trabalham.

A execução desta etapa da pesquisa restou prejudicada em razão da baixa adesão obtida dos Tribunais de Contas à pesquisa. Tal fato ocasionou a recepção de poucas respostas por Tribunal, o que prejudica o cálculo da pontuação dos Tribunais de Contas, já que esta tomaria por base a percepção de poucos técnicos da instituição.

Não obstante, objetivando ilustrar o procedimento de cálculo, será utilizada escala obtida em um Tribunal de Contas com vistas a obter o seu escore. O objetivo é quantificar, com base na escala construída, o escore para representar o resultado da aplicação da escala no referido Tribunal de Contas.

A quantidade de respondentes neste TC é de 5 pessoas. O escore do TC será a média dos escores obtidos por cada um dos técnicos.

O referido TC teve um escore de **47,20** (vide o Apêndice 5). Este escore é alto ou baixo?

A escala possui 18 itens, então cada respondente pode ir de um mínimo de 18 pontos até um

máximo de 90 pontos. O valor neutro da escala implica uma pontuação de 54. O escore do TC em questão indica um nível de maturidade satisfatório.

Naturalmente que quanto mais técnicos responderem à escala, mais confiável será o escore obtido para o Tribunal.

A mesma análise será feita para um outro TC que apresentou 19 respostas.

Este outro TC apresentou um escore maior, no valor de **51,53**.

Mas por que o 2º TC teve um escore maior que o primeiro? O que o primeiro TC deve fazer para melhorar?

Calculando-se o escore de cada um dos TCs em cada fator, ter-se-á uma indicação de onde o TC deve melhorar. Exemplo:

Tribunal	F1	F2	F3	F4	F5	F6	F7	TOTAL
01	13,58	15,32	7,53	1,79	8,84	2,16	2,32	51,53
02	11,40	15,60	7,20	2,20	6,60	2,40	1,80	47,20

A aplicação da escala aos demais Tribunais forneceria uma relação com a pontuação de cada Tribunal, com indicação de onde poderiam melhorar.

3 - RESULTADOS/CONCLUSÕES

Diferentemente das abordagens mais comuns para o desenvolvimento de escalas de maturidade, esta pesquisa oferece uma metodologia nova, baseada na construção de uma escala de Likert.

Acredita-se que esta abordagem tenha a vantagem de produzir um escore de maturidade analítica para o Tribunal de Contas a partir das percepções dos técnicos quanto a realidade existente. A metodologia também permite que se identifique em quais áreas o TC deve investir para melhorar o seu nível de maturidade analítica.

A escala final proposta se compõe dos seguintes itens:

Fator 1 - Existência de pessoal capacitado em softwares e técnicas analíticas

Q04 - As iniciativas de uso de análise de dados são isoladas

Q06 - O Tribunal utiliza *dashboard* para monitorar métricas chaves para o controle

Q19 - O uso de softwares que suportam análise de dados é difundido entre os auditores

Q21 - O uso de mineração de texto para auxiliar no exercício do controle externo ainda não é uma realidade

Q22 - O Tribunal dispõe de técnicos capacitados em softwares analíticos

Fator 2 - Existência e uso de dados

Q07 - Além dos dados obtidos com os jurisdicionados, o Tribunal também disponibiliza aos técnicos dados obtidos de outras instituições

Q08 - O Tribunal possui um setor que tem como missão gerenciar informações para dar suporte às atividades de controle externo do Tribunal

Q10 - A análise de dados é utilizada para auxiliar na detecção de fraudes

Q11 - O Tribunal possui em sua estrutura organizacional um setor encarregado de garantir a qualidade dos dados utilizados para a realização das auditorias/fiscalizações

Fator 3 - Apoio institucional para o uso de análise de dados

Q03 - O Tribunal oferece a seus técnicos capacitação em análise de dados

Q05 - Inexiste incentivo por parte do Tribunal para que os técnicos se capacitem em análise de dados

Q12 - Inexiste apoio da alta administração do Tribunal ao uso de análise de dados para realizar a atividade de controle

Fator 4 - Infraestrutura inadequada para manipular grandes bases de dados

Q13 - Os recursos computacionais (hardware) disponibilizados aos técnicos são inadequados para trabalhar com grandes bases de dados

Fator 5 - Vontade institucional para o uso de análise de dados

Q09 - A análise de risco é feita de forma subjetiva, valendo-se da experiência do técnico

Q18 - O Tribunal possui ações estratégicas previstas explicitamente em seu Planejamento Estratégico

visando aumentar sua capacidade analítica

Q20 - Os técnicos têm dificuldade de dispor dos dados necessários à realização de auditorias/fiscalizações

Fator 6 - Competência técnica para manuseio de grandes bases de dados

Q15 - Os técnicos não se sentem confortáveis para trabalhar com grandes bases de dados

Fator 7 - Tomada de decisões baseada em fatos

Q16 - A seleção de objetos de fiscalização é realizada com base na experiência dos técnicos, sem o auxílio de análise de risco

Em razão da baixa adesão dos Tribunais pesquisados, não foi possível uma mensuração mais robusta dos escores, dado o baixo número de respondentes em cada Tribunal. Ainda assim, apresenta-se abaixo um quadro com os escores obtidos pelos Tribunais que atenderam à pesquisa e tiveram 5 ou mais respondentes.

Tabela de escores dos Tribunais

Tribunal	F1	F2	F3	F4	F5	F6	F7	TOTAL	QTD RESP.
01	13,58	15,32	7,53	1,79	8,84	2,16	2,32	51,53	19
02	11,40	15,60	7,20	2,20	6,60	2,40	1,80	47,20	5
03	10,43	14,43	7,00	2,43	8,71	2,57	2,00	47,57	7
04	9,71	12,86	6,71	1,86	7,29	2,00	2,29	42,71	7

Como já dito, a pouca adesão dos tribunais à pesquisa, faz com que os resultados obtidos para os tribunais não sejam robustos o suficiente para permitir uma avaliação definitiva.

Não obstante, os dados colhidos na amostra piloto permitiram a elaboração da escala de maturidade analítica que poderá, efetivamente, ser utilizada pelos tribunais com vistas a avaliar seu grau de maturidade analítica e também, em que áreas será possível melhorar.

Como proposta de melhoria, pretende-se, a exemplo do que foi feito no artigo “Uma escala para medir a infraestrutura escolar” (Neto et al., 2013), utilizar a Teoria da Resposta ao Item para criar níveis de maturidade nos quais os Tribunais poderão ser classificados.

4 - REFERÊNCIAS

Davenport, Thomas H., Harris; Jeanne G.; - Competição Analítica: vencendo através da nova ciência: como decisões baseadas em fatos e business intelligence transformam os negócios - Rio de Janeiro, Elsevier 2007

Sales, Leonardo Jorge - Proposta de Modelo de classificação do risco de contratos públicos; Brasília, 2015 - Dissertação de Mestrado.

Carvalho, Ricardo Silva - Modelos Preditivos para Avaliação de Risco de Corrupção de Servidores Públicos Federais; Brasília, 2015 - Dissertação de Mestrado

Decisão do TCU - TC 010.247/2015-7 Acórdão 539/2016.

Remis Balaniuk - A Mineração de Dados como apoio ao Controle Externo, Revista do TCU

Miklos A. Vasarhelyi e Siripan Kuenkaikaew - The Predictive Audit Framework - The International Journal of Digital Accounting Research, Vol. 13, 2013, pp. 37-71

José Luiz Kugler - Competência Analítica: Conceitos e Estratégias para a construção da empresa inteligente - São Paulo, Saraiva, 2013

Laila Costa Damasceno; Inteligência Analítica: um estudo de caso de maturidade analítica - Brasília 2012 Monografia (bacharelado)

Thomas H. Davenport e Sirkka L. Jarvenpaa; Strategic Use of Analytics in Government - Managing for Performance and Results Series 2008 IBM Center for The Business of Government

Clayton Ferreira da Silva, Erick Muzart Fonseca dos Santos, Mônica Cotrim Chaves, Wesley Vaz e Remis Balaniuk - Dados abertos: uma estratégia para o aumento da transparência e modernização da gestão pública; Revista do TCU

Carvalho, Ricardo Silva - Modelos Preditivos para a Avaliação de Risco de Corrupção de Servidores Públicos Federais - Brasília 2015 (Dissertação de Mestrado)

Gledson Pompeu - Contribuições da Auditoria Contínua para a Efetividade do Controle Externo - Brasília 2012

Gledson Pompeu Corrêa da Costa, Tiago Alves de Couveia Lins Dutra; Auditoria financeira na do Big Data: nova possibilidades para avaliação e resposta a riscos em demonstrações financeiras do Governo Federal - Revista do TCU

Hair Jr., Joseph F.; Babin, Barry; Money, Arthur H.; Samouel, Phillip - Fundamentos de

Métodos de Pesquisa em Administração – Porto Alegre, Bookman 2005.

Vieria, Valter Afonso; Escalas em marketing: métricas de resposta do consumidor e de desempenho empresarial – São Paulo, Atlas 2011

Vergara, Sylvia C.; Projetos e Relatórios de Pesquisa em Administração, 15 Ed. – São Paulo, Atlas 2014

Johns, Rob; Likert Items and Scales – Survey Question Bank: Methods Fact Sheet 1 (March 2010)

Elliot, Ligia Gomes (org.); Instrumentos de Avaliação e Pesquisa: caminhos para construção e validação – Wak editor – 2012

SAS Institute Inc.; How to compete on analytics: the analytical center of excellence – white paper

Infocontas – Rede Nacional de Informações Estratégicas para o Controle Externo: criação e regulamentação - 2013

5 - APÊNDICES

5.1 Apêndice 1 – Cronograma previsto e executado

Atividades	Status	Ago	Set	Out	Nov	Dez	Jan	Fev	Mar	Abr	Mai	Jun	Jul	Ago
Pesquisa bibliográfica e revisão da literatura	P	X	X	X	X	X								
	R	X	X	X	X	X								
Levantamento de contatos junto aos TCs	P	X	X	X										
	R	X	X	X	X	X	X	X	X	X				
Elaboração do instrumento de coleta	P													
	R				X	X	X							
Validação do instrumento de coleta junto a especialistas	P				X	X	X							
	R					X								
Aplicação de questionário piloto	P						X	X	X					
	R												X	X
Avaliação do questionário e ajustes	P								X					
	R													X
Aplicação do questionário final	P									X	X			
	R													
Análise dos resultados	P											X	X	
	R													
Construção da escala de maturidade analítica	P												X	X
	R											X	X	X
Redação do relatório final	P												X	X
	R													X

P – Previsto; R - Realizado

5.2 Apêndice 2 – Dificuldades encontradas

Durante a execução da pesquisa duas situações dificultaram a realização da pesquisa. A primeira dificuldade deve-se à disponibilização pelo TCE-RJ de um aplicativo para aplicação de questionários *on-line*.

No âmbito da Coordenadores de Auditorias Temáticas e Operacionais – CTO estava em uso o aplicativo SurveyMonkey, com base no qual foram realizadas algumas pesquisas no âmbito da Coordenadoria, bem como uma das pesquisas realizada na edição de 2015 do projeto de pesquisas da ECG. Ocorre que a subscrição do serviço, paga pelos técnicos da área de TI da CTO, cujo prazo era de um ano, findou durante a realização da pesquisa. Foi feito encaminhamento à Secretaria-Geral de Controle Externo para aquisição de nova subscrição de aplicativo semelhante que acabou por resultar na disponibilização, após algum tempo, pela Diretoria-Geral de Informática – DGI do aplicativo LimeSurvey, para cuja utilização não se tinha experiência anterior.

Outra situação que merece destaque é a baixa adesão dos técnicos dos Tribunais de Contas à pesquisa realizada. Este fato comprometeu o cálculo de escores para os Tribunais.

5.3 Apêndice 3 – Formulário para validação dos itens por especialistas

1-INTRODUÇÃO:

O Tribunal de Contas do Estado do Rio de Janeiro – TCE-RJ, por intermédio de sua Escola de Contas e Gestão – ECG, mantém um Programa de Pesquisa instituído pela Resolução ECG/TCE RJ nº 8 de 19/08/2014.

Para a edição de 2016 nossa pesquisa, aprovada pela ECG, denomina-se: “**Controle Externo Baseado em Análise de Dados: estão os Tribunais de Contas preparados?**”.

A pesquisa objetiva endereçar o problema de se avaliar em que medida o nível de maturidade analítica existente nos Tribunais de Contas permite a efetiva implementação de uma estratégia de controle que se baseie em análise de dados.

Para tanto estamos propondo a criação uma escala de maturidade analítica com base na qual os Tribunais de Contas possam ser avaliados. Isso exige que se se mensure um conceito abstrato, qual seja, o conceito de “Capacidade Analítica do Tribunal de Contas”.

Davenport e Harris no livro [Competição Analítica – vencendo através da nova ciência](#) propõem que este conceito possa ser mensurado a partir de 3 fatores: fatores organizacionais, fatores humanos e fatores tecnológicos.

Essa referência, juntamente com as Portarias TCU [141 de 01/04/15](#) e [130 de 27/03/15](#), foram o ponto de partida para a estruturação de uma proposta inicial itens (variáveis) que comporão uma escala de likert para a mensuração do conceito em referência, bem como de um questionário não estruturado que tem por objetivo trazer para a pesquisa o conhecimento de especialistas que irá contribuir na definição final tanto dos itens quanto da escala a ser efetivamente utilizada.

Esperamos, **com isso**, elaborar um instrumento de mensuração, bem como uma escala, que sejam mais aderentes às particularidades de instituições de controle públicas, em especial os Tribunais de Contas.

Para mais informações sobre o Projeto de Pesquisa mantido pela ECG/TCE-RJ, consultar a página <http://www.ecg.tce.rj.gov.br/linhas-de-pesquisa>.

2-OBJETIVO DESTA CONSULTA

Esta consulta, endereçada a especialistas no uso de análise de dados em instituições de controle, tem por objetivo colher subsídios que nos permitam identificar quais elementos são percebidos como fundamentais para a elaboração de um instrumento de coleta de dados que permita colher informações junto aos Tribunais de Contas com vistas a mensurar a capacidade analítica dos mesmos.

Quaisquer dúvidas, esclarecimentos ou comentários adicionais acerca dos objetivos da pesquisa ou das questões propostas devem ser endereçadas aos pesquisadores:

Marcos Ferreira da Silva.

Tel. (21) 3231-4726 | (21) 99193-4117,

email: marcosfs@tce.rj.gov.br

Bruno Mattos Souza de Souza Melo.

Tel. (21) 3231-4726 | (21) 97914-8321

email: brunosm@tce.rj.gov.br

PARTE A – ITENS PARA MENSURAÇÃO DO CONCEITO DE CAPACIDADE/COMPETÊNCIA ANALÍTICA

Apresentamos a seguir, com vistas a críticas e sugestões, um conjunto de itens que foram elaborados com base na literatura consultada.

Não obstante, entendemos que um conjunto final de itens, **que** tenham por objetivo mensurar a capacidade analítica dos Tribunais de Contas, não poderia prescindir de uma apreciação e de críticas por parte de pessoas com experiência na adoção e/ou implementação de técnicas de análise de dados em instituições de controle.

Apresentamos, **para críticas e sugestões**, os itens a seguir elencados que buscam mensurar o conceito de competência analítica a partir de fatores organizacionais, humanos e tecnológicos, seguindo a proposta de *Harris e Davenport*.

	Item	Fator
1	O Tribunal oferece a seus técnicos capacitação em análise de dados avançada (regressão múltipla, <i>deep learning</i> , regras de associação, máquinas de vetor suporte, classificador bayesiano etc.).	Organizacional (capacitação)
2	O Tribunal utiliza metodologia formalmente instituída para selecionar objetos de fiscalização com base em análise de risco.	Organizacional (processos)
3	O uso de softwares para análise avançada de dados (SAS, R, Stata Python etc.) é bastante difundido entre os auditores.	Pessoal (habilidades)
4	Os técnicos que realizam auditoria se sentem confortáveis para trabalhar com grandes volumes de dados.	Pessoal (habilidades)
5	O Tribunal possui em sua estrutura organizacional um setor encarregado de garantir a qualidade e a confiabilidade dos dados utilizados para o planejamento das auditorias.	Organizacional (estrutura)
6	O Tribunal possui ações estratégicas previstas explicitamente em seu Planejamento Estratégico visando aumentar sua capacidade analítica.	Organizacional (cultura.)
7	A alta administração do Tribunal apoia o uso de análise de dados para planejar e executar as auditorias	Organizacional (patrocínio.)
8	O Tribunal possui uma cultura de tomada de decisões com base em fatos.	Organizacional (cultura)

9	A definição ou priorização das auditorias tem por base critérios objetivos de mensuração de risco.	Organizacional (processos)
10	Nos dois últimos anos o Tribunal realizou pelo menos uma auditoria/fiscalização em que utilizou técnicas de mineração de dados tais como análise de agrupamentos, regressão logística, mineração de texto, regras de associação, máquina de vetor suporte etc.	Pessoal (habilidade)
11	O Tribunal utiliza recursos de geoprocessamento para auxiliar no planejamento ou execução de auditorias.	Tecnologia (software)
12	O Tribunal utiliza <i>dashboard</i> para monitorar métricas chaves em ao menos uma área de fiscalização (educação, saúde, previdência, obras, contratações etc.)	Tecnologia (software)
13	O Tribunal utiliza técnicas para análise automática do conteúdo de informações não estruturadas (texto) com vistas a gerar subsídios para o exercício do controle.	Organizacional (processos)
14	O Tribunal utiliza técnicas quantitativas preditivas para subsidiar a classificação de risco.	Organizacional (processos)
15	Os técnicos do Tribunal utilizam técnicas de mineração de dados para auxiliar na detecção de fraudes.	Organizacional (processos)
16	O Tribunal dispõe de técnicos capacitados em técnicas de análise de dados e em softwares analíticos em todos os setores do controle externo que realizam auditorias e fiscalizações.	Pessoal (habilitação)
17	Os dados necessários ao planejamento das fiscalizações estão disponíveis para uso por todos os técnicos, independentemente de sua área de atuação.	Tecnologia (dados)
18	O Tribunal dispõe de um setor que tem como missão gerenciar informações para dar suporte às atividades de fiscalização do Tribunal, pensando soluções tecnológicas para o controle e obtendo e processando informações que vêm de fora do Tribunal.	Organizacional (processos)
19	O Tribunal incentiva os técnicos a buscarem capacitação em análise de dados.	Organizacional (capacitação)
20	O Tribunal possui recursos computacionais (<i>hardware</i>) que permitem trabalhar com grandes volumes de dados	Tecnologia (hardware)
21	Além dos dados colhidos junto aos jurisdicionados, o Tribunal também disponibiliza aos técnicos dados obtidos de outras instituições.	Tecnologia (dados)
22	Iniciativas de uso de análise de dados são localizadas em apenas alguns setores do controle externo por contarem com o apoio apenas da gerência local.	Organizacional (patrocínio)

1 - A seu ver, algum aspecto relativo à mensuração da capacidade analítica dos Tribunais não foi contemplado nos itens acima? Caso algum aspecto não tenha sido contemplado, por favor nos informe quais seriam.

2 – A seu ver, a mensuração da capacidade analítica a partir de fatores organizacionais, humanos e tecnológicos é adequada? Caso entenda que não, por favor nos informe as razões.

PARTE B – ESCALA DE MATURIDADE

Davenport e Harris propõem uma escala de maturidade com 5 níveis:

Nível 1 – Neste estágio, a entidade ainda não possui os pré-requisitos para a inteligência analítica, ou seja, não possuem dados consistentes e de alta qualidade para a tomada de decisões.

Nível 2 – Neste estágio, a entidade possui iniciativas isoladas de análise de dados, o patrocínio limita-se ao nível gerencial.

Nível 3 – Neste estágio, a entidade possui o patrocínio da alta administração. Há o incentivo aberto à adoção de uma cultura mais baseada em fatos.

Nível 4 – Neste estágio, a entidade busca o desenvolvimento de competências analíticas de primeira linha na entidade como um todo. Implementa-se os planos desenvolvidos no estágio anterior.

Nível 5 – Neste estágio, a inteligência analítica está disseminada por toda a instituição, constituindo a chave para a sua estratégia e vantagem competitiva. A entidade colhe continuamente os benefícios da competência analítica desenvolvida no âmbito de toda a empresa e se concentra na renovação analítica contínua.

A nossa escala resultará da análise dos questionários encaminhados aos Tribunais com uso de técnicas de análise multivariada (análise fatorial, análise de agrupamento e teoria da resposta ao item).

PARTE C – QUESTIONÁRIO ABERTO

Objetivando colher subsídios para a consecução da pesquisa, gostaríamos de colher a sua opinião sobre alguns aspectos relacionados à competência analítica de uma instituição de controle. Para tanto, pedimos que responda às questões a seguir:

C.1 - No âmbito de uma instituição de controle, o que seria (ou definiria), a seu ver, uma “**Estratégia de Controle Baseada em Análise de Dados**”? (Exponha seu entendimento sobre este tópico.)

C.2 – Em sua opinião, quais seriam os principais entraves à implantação de uma cultura analítica em uma instituição de controle?

C.3 – Em sua opinião, o que deve ser aperfeiçoado ou implementado nos Tribunais de Contas para que tais instituições possam realizar de forma efetiva a auditoria preditiva? [\[Ref 1\]](#)[\[Ref 2\]](#) (Vide referências contidas nos links para entendimento do que estamos considerando que seja auditoria preditiva)

C.4 – Fique à vontade para expor ideias complementares que não foram abordadas, **mas** que você considera

importantes para uma proposta de mensuração da competência analítica dos Tribunais de Contas.

C.5- Havendo necessidade de esclarecer alguma resposta da entrevista como poderemos contactá-lo?

Nome:

Instituição:

Telefone

E-mail:

5.4 Apêndice 4 – Quantidade de Tribunais participantes da pesquisa e quantidade de respostas obtidas

- Quantidade de Tribunais: 17
- Quantidade de Respostas: 105 respondentes
- Média por Tribunal: 6,17

Estes dados indicam a baixa adesão dos TCs à pesquisa.

5.5 Apêndice 5 – Análise exploratória dos dados da amostra-piloto

Importação e preparo das bases de dados

Os comandos a seguir têm por objetivo realizar a importação das bases de dados e realizar seu pré-processamento, antes que a análise dos dados propriamente dita possa ser realizada.

```
library(readxl)
library(likert)
library(psych)
library(tidyr)
library(ggplot2)
library(corrplot)

# Importa a base de dados
setwd("F:\\Analise_Questionario_Pesquisa")

dados1 <- read_excel("results-survey273734_DtGer-11-08-2017.xlsx") # piloto
dados1 <- dados1[,c(9:31)]

dados2 <- read_excel("results-survey913162_DtGer-25-05-2017.xlsx") # Link gratuito - dados sequestrados
dados2 <- dados2[,c(5:27)]

dados3 <- read_excel("results-survey266345 (1).xlsx") # 2o Link gratuito
dados3 <- dados3[,c(5:27)]
dados3 <- subset(dados3, !is.na(dados3$`Você faz parte de qual Tribunal de Contas?`))

dados4 <- read_excel("results-survey376929-DtGer29-08-2017.xlsx")
dados4 <- dados4[,c(5:27)]
dados4 <- subset(dados4, !is.na(dados4$`Você faz parte de qual Tribunal de Contas?`))

dados <- rbind(dados1, dados2, dados3, dados4)
rm(dados1, dados2, dados3, dados4)

# Limpeza e organização dos dados
dados <- as.data.frame(dados)
#dados <- dados[, -(1:8)]
dados <- dados[-(1:9), ] # Exclui registros de teste: 9 primeiras linhas
names(dados) <- c("TC", paste(rep("Q", 22), sprintf("%02.0f", 1:22), sep=""))#
Renomeia as variáveis

O <- c("Q03", "Q04", "Q05", "Q08", "Q09", "Q10", "Q11", "Q12", "Q16", "Q17", "Q18", "Q21")# Agrupa as questões
H <- c("Q01", "Q02", "Q15", "Q19", "Q22")
```

```
T <- c("Q06", "Q07", "Q13", "Q14", "Q20")

#names(dados)[-1] <- ifelse(names(dados)[-1] %in% 0, paste(names(dados)[-1], "0", sep = "-"),
#                             ifelse(names(dados)[-1] %in% H, paste(names(dados)[-1], "H", sep = "-"),
#                                     paste(names(dados)[-1], "T", sep = "-")))

dados <- dados[complete.cases(dados),]

# Converter string de resposta para código
troca_troca <- function(x){
  x[x == "DISCORDO COMPLETAMENTE"] <- "1"
  x[x == "DISCORDO"] <- "2"
  x[x == "NAO CONCORDO E NEM DISCORDO"] <- "3"
  x[x == "CONCORDO"] <- "4"
  x[x == "CONCORDO COMPLETAMENTE"] <- "5"
  as.integer(x)
}

dados.int <- dados
dados.int[,-1] <- as.data.frame(sapply(dados[,-1], troca_troca))

### Construir data frame para usar o pacote likert
###-----
# Solução do autor do pacote
dados.lik <- dados

niveis <- sapply(dados.lik[,-1], unique)
niveis <- Reduce(union, niveis)
niveis <- niveis[c(5, 2, 4, 3, 1)]

# Conversão dos itens para fator.
dados.lik[,-1] <- as.data.frame(apply(dados.lik[,-1], 2, factor))

# Inserção de niveis faltantes em alguns itens
for(i in names(dados.lik[,-1])) { # substituir por apply?
  dados.lik[,i] <- factor(dados.lik[,i], levels=niveis)
}
```

Quantidade de Respostas por Tribunal de Contas Pesquisado

Após a aplicação do questionário-piloto, foram obtidas respostas de 17 Tribunais de Contas.

Análise dos Dados do TCE-RJ

Para realizar a análise dos dados relativos ao TCE-RJ, é necessário selecionar os registros de interesse, o que é feito a seguir:

As estatísticas descritivas para os itens são apresentadas a seguir. Antes, contudo, faz-se necessário realizar a alteração nos valores das questões que foram balanceadas:

```
# Questões balanceadas: "Q04", "Q05", "Q09", "Q12", "Q13", "Q14", "Q15", "Q16",
"Q20", "Q21"
reverte <- function(x) {
  x[x == 1] <- 5
  x[x == 2] <- 4
  x[x == 4] <- 2
  x[x == 5] <- 1
  as.integer(x)
}
questoes_balanceadas <- c("Q04", "Q05", "Q09", "Q12", "Q13", "Q14", "Q15", "Q16",
", "Q20", "Q21")

tce.rj.int_rev <- tce.rj.int
tce.rj.int_rev[, questoes_balanceadas] <- lapply(tce.rj.int_rev[, questoes_bala
nceadas], reverte)
```

Feito o tratamento dos dados, serão feitas algumas análises descritivas.

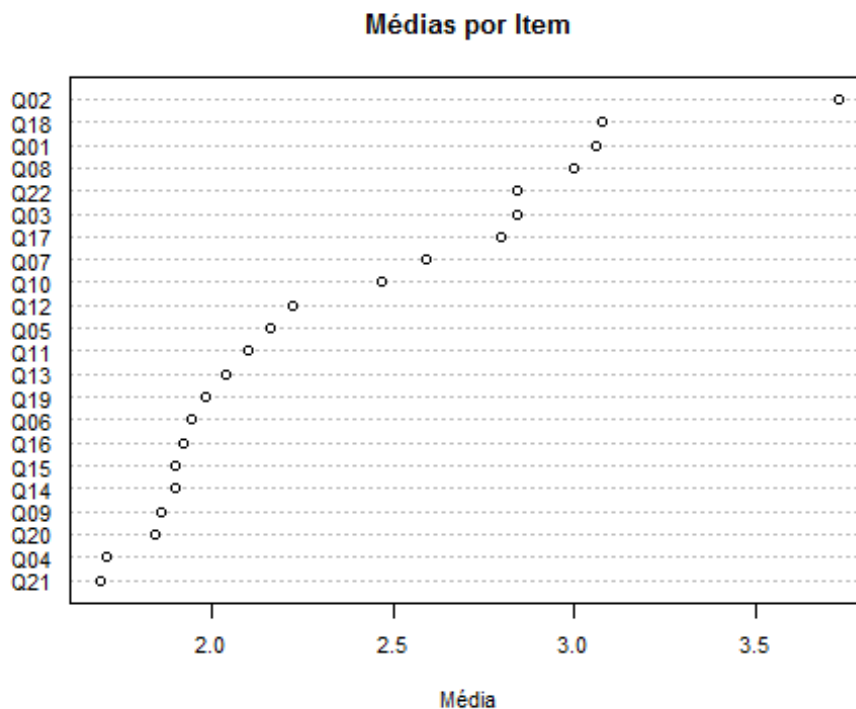
```
# Média, desvio padrão e quantidade de questionários
estat <- as.data.frame(t(sapply(tce.rj.int_rev, function(x) c(Media=round(mean(
x), 2), Soma=sum(x), DP=round(sd(x), 2), Qtd_Quest=length(x))))))
```

estat

##	Media	Soma	DP	Qtd_Quest
## Q01	3.06	150	1.01	49
## Q02	3.73	183	0.88	49
## Q03	2.84	139	1.09	49
## Q04	1.71	84	0.50	49
## Q05	2.16	106	0.55	49
## Q06	1.94	95	0.75	49
## Q07	2.59	127	1.02	49
## Q08	3.00	147	1.12	49
## Q09	1.86	91	0.54	49
## Q10	2.47	121	1.02	49
## Q11	2.10	103	0.94	49
## Q12	2.22	109	0.51	49

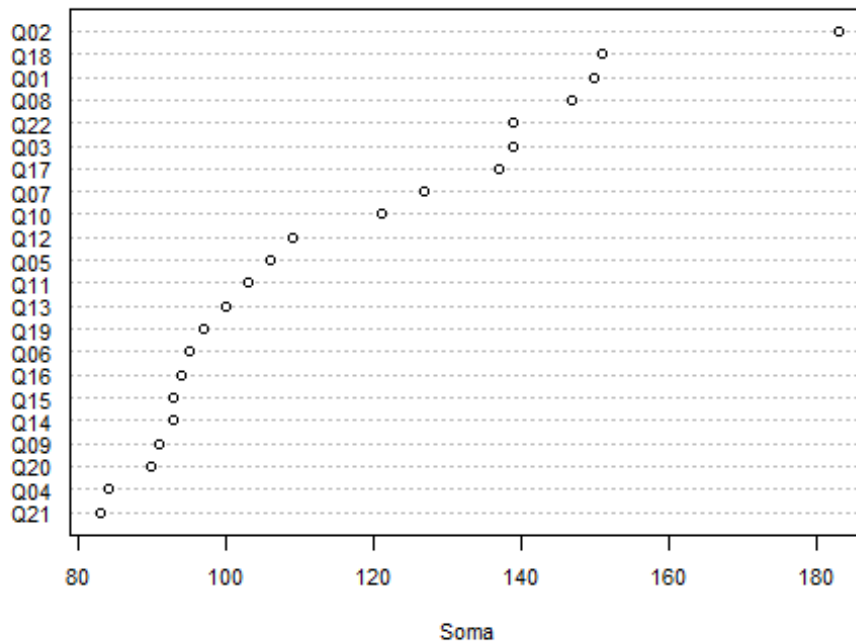

```
## Q13  2.04  100 0.64      49
## Q14  1.90   93 0.62      49
## Q15  1.90   93 0.62      49
## Q16  1.92   94 0.57      49
## Q17  2.80  137 0.98      49
## Q18  3.08  151 0.81      49
## Q19  1.98   97 0.90      49
## Q20  1.84   90 0.55      49
## Q21  1.69   83 0.58      49
## Q22  2.84  139 1.03      49
```

```
estat <- estat[order(estat$Media),]
dotchart(estat$Media, labels = row.names(estat), cex=0.7, main="Médias por Item", xlab = "Média")
```



```
estat <- estat[order(estat$Soma),]
dotchart(estat$Soma, labels = row.names(estat), cex=0.7, main="Escore Total por Item", xlab="Soma")
```

Score Total por Item



Calcular a frequência de cada escore por questão.

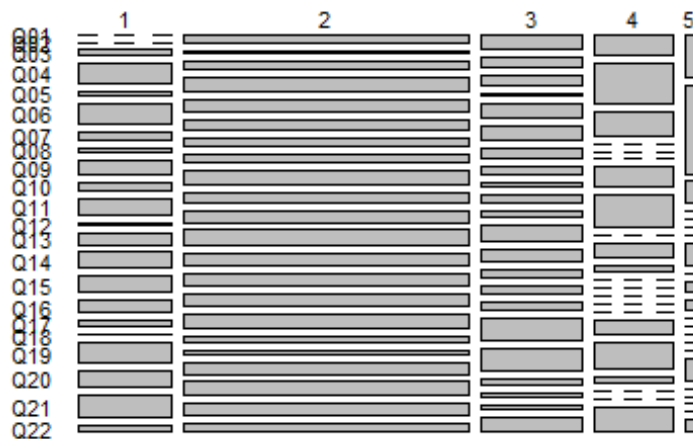
```
for(i in names(tce.rj.int_rev)) {
  tce.rj.int_rev[,i] <- factor(tce.rj.int_rev[,i], levels=as.character(1:5))
}
```

```
## reshape para calcular a frequencia em cada escore por item
knitr::kable(t(sapply(tce.rj.int_rev, table)))
```

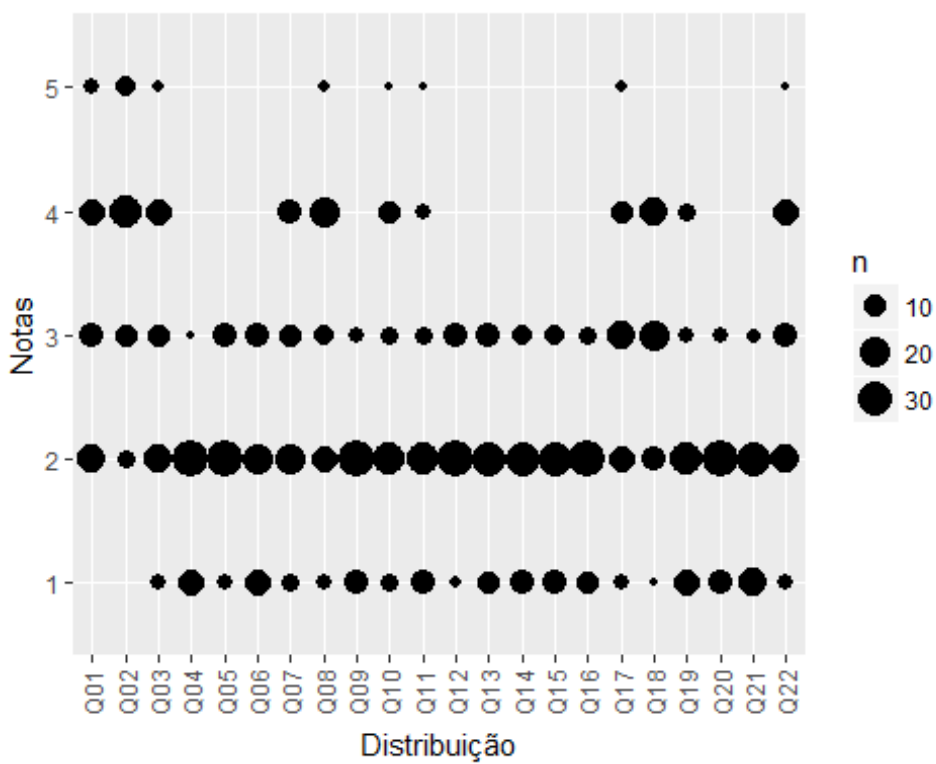
	1	2	3	4	5
Q01	0	19	12	14	4
Q02	0	6	9	26	8
Q03	4	19	9	15	2
Q04	15	33	1	0	0
Q05	4	33	12	0	0
Q06	15	22	12	0	0
Q07	6	21	9	13	0
Q08	4	16	7	20	2
Q09	11	34	4	0	0
Q10	6	26	6	10	1
Q11	12	26	6	4	1

Q12	2	34	13	0	0
Q13	9	29	11	0	0
Q14	12	30	7	0	0
Q15	12	30	7	0	0
Q16	10	33	6	0	0
Q17	4	15	19	9	2
Q18	1	11	20	17	0
Q19	15	25	4	5	0
Q20	12	33	4	0	0
Q21	18	28	3	0	0
Q22	4	17	12	15	1

```
mosaicplot(sapply(tce.rj.int_rev, table), las=1, main="")
```



Objetivando obter visualizações da distribuição das respostas no questionário, apresenta-se um conjunto de visualizações de caráter exploratório, que objetiva a obtenção de uma melhor compreensão dos dados.

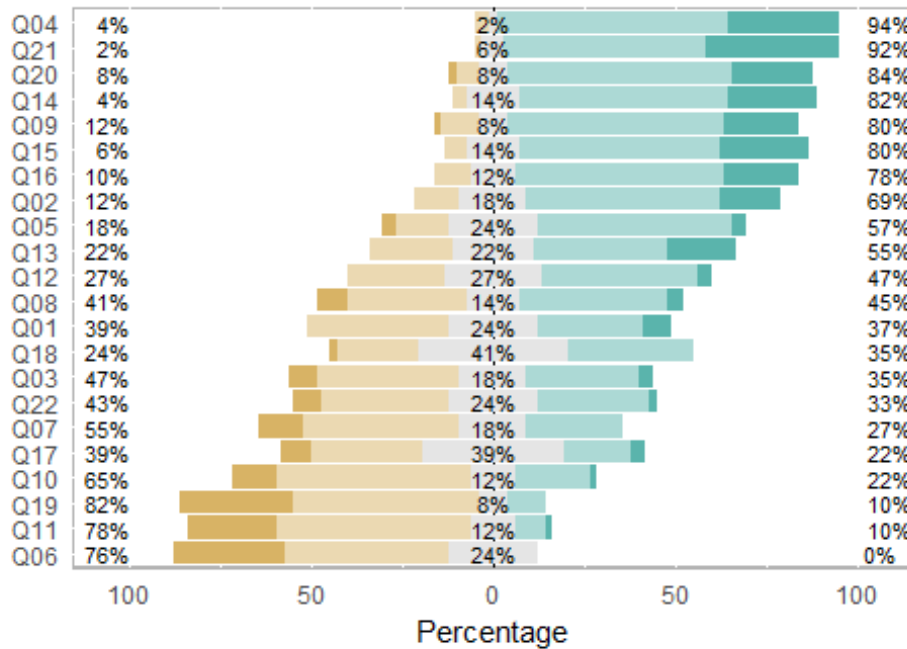


Conversão de fatores para valores inteiros novamente.

```
tce.rj.int_rev[] <- lapply(tce.rj.int_rev, function(x) as.integer(as.character(x)))
```

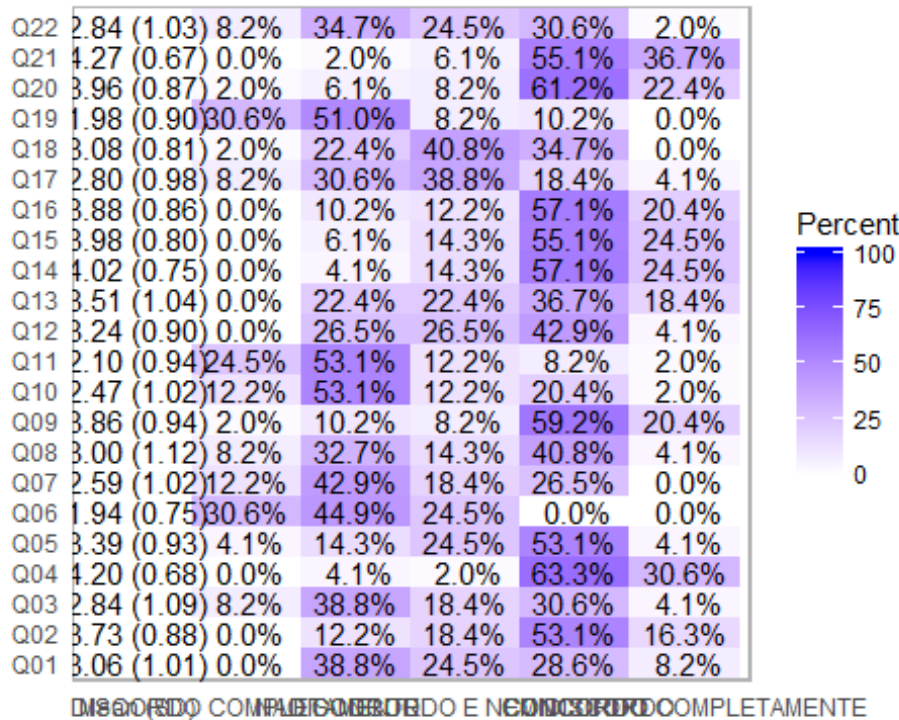
Gráfico com todas as questões

```
plot(likert(tce.rj.lik))
```



IPLETAMENTE DISCORDO NAO CONCORDO E NEM DISCORDO CONCC

```
plot(likert(tce.rj.lik), type = "heat")
```

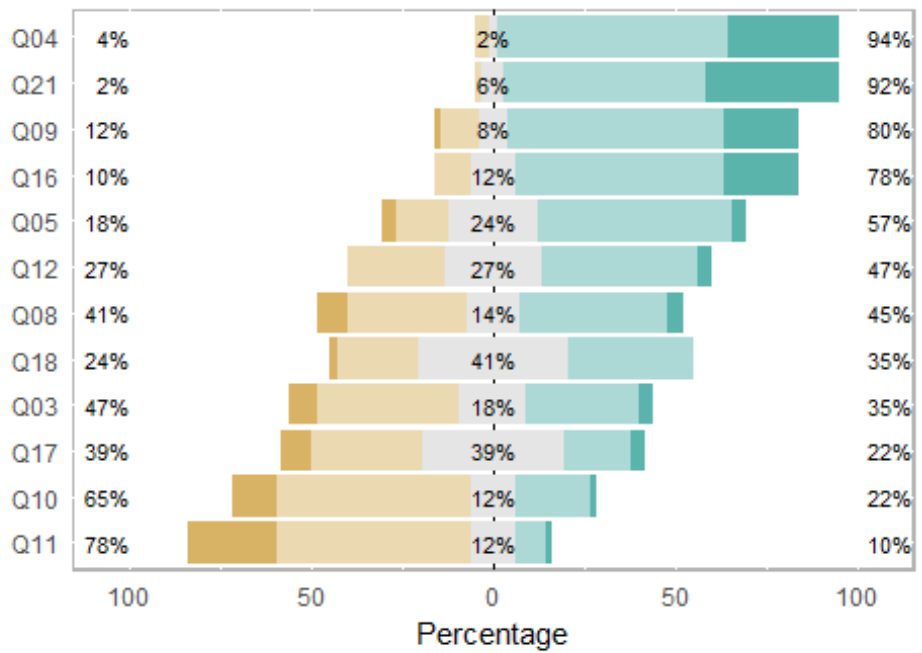


Para auxiliar na análise, as questões revertidas foram: Q04, Q05, Q09, Q12, Q13, Q14, Q15, Q16, Q20, Q21. O gráfico evidencia os percentuais de respondentes que adotaram uma atitude neutra, e os percentuais de

discordância e concordância. Assim, por exemplo, para a questão Q04 (As iniciativas de análise de dados são isoladas), 4% discordaram da assertiva e 94% concordaram com a assertiva.

Gráfico com as questões do fator organizacional:

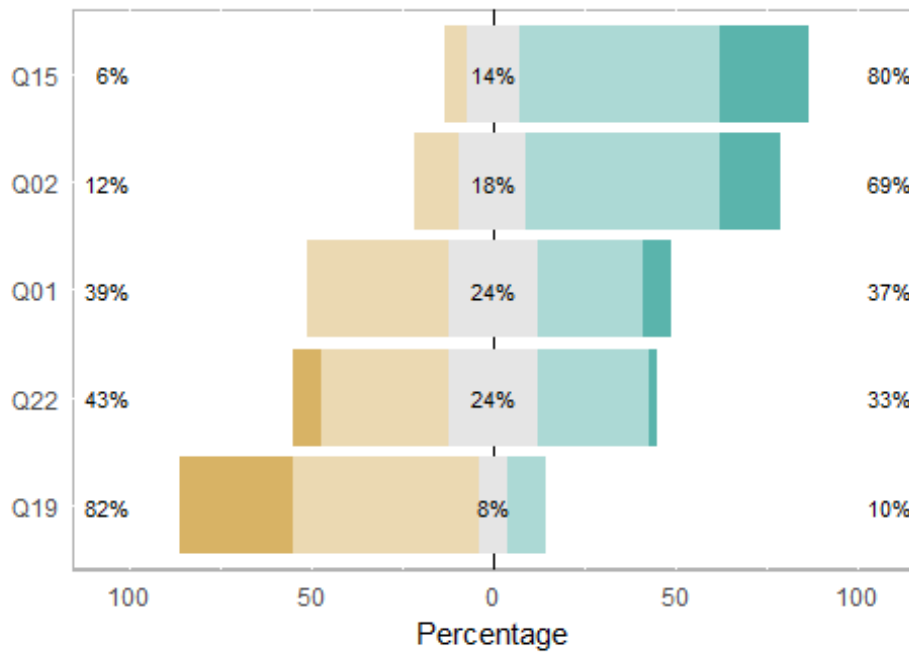
```
#invisible(likert.options(text.size=0.5))
plot(likert(tce.rj.lik[, 0]))
```



IPLETAMENTE DISCORDO NAO CONCORDO E NEM DISCORDO CONCC

Gráfico com as questões do fator humano:

```
plot(likert(tce.rj.lik[, H]))
```

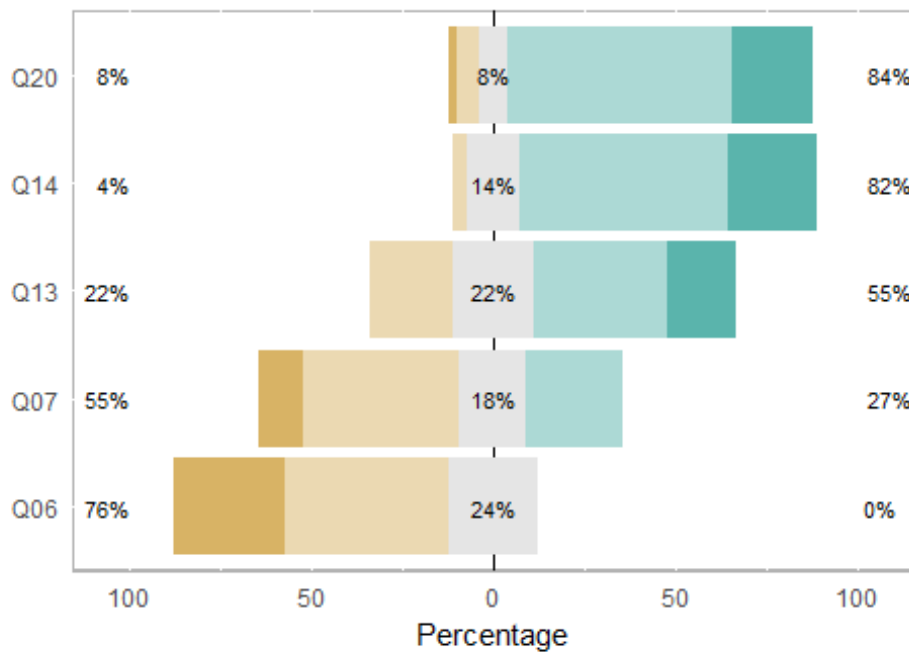


IPLETAMENTE DISCORDO NAO CONCORDO E NEM DISCORDO CONCC

As questões revertidas são: Q15

Gráfico com as questões do fator tecnológico:

`plot(likert(tce.rj.lik[, T]))`



IPLETAMENTE DISCORDO NAO CONCORDO E NEM DISCORDO CONCC

As questões revertidas são: Q13, Q14, Q20

Análise Fatorial Exploratória

De acordo com Hair Jr. et al. (2005) *apud* Zanella et al. (2010), "a análise fatorial é um método estatístico multivariado cujo propósito principal é definir a estrutura subjacente em uma matriz de dados. Essa técnica aborda o problema de analisar a estrutura das inter-relações entre um grande número de variáveis, definindo um conjunto de dimensões latentes chamadas de fatores. Primeiramente, é possível identificar as dimensões e então determinar o grau em que cada variável é explicada por cada dimensão".

Na construção de escalas de atitude, o uso da análise fatorial tem por objetivo identificar as variáveis latentes e auxiliar na escolha dos itens que irão compor a escala.

Considerações sobre o tamanho da amostra:

A análise fatorial dos dados colhidos exige que se disponha de um conjunto razoável de observações. Hair et al. (2009) recomendam que a quantidade de observações seja superior a 50 casos, sendo aconselhável no mínimo 100 observações. A relação entre observações e número de variáveis recomendada por estes autores é de 5:1, sendo, no entanto, desejável 10:1.

Na presente pesquisa, foi obtida uma amostra de 49 observações, fato que constitui uma limitação à construção de uma escala confiável.

Análise da estrutura de correlações:

A matriz de correlações é apresentada a seguir:

Matriz de correlações

```
cortcerj <- cor(tce.rj.int)
round(cortcerj, 2)
```

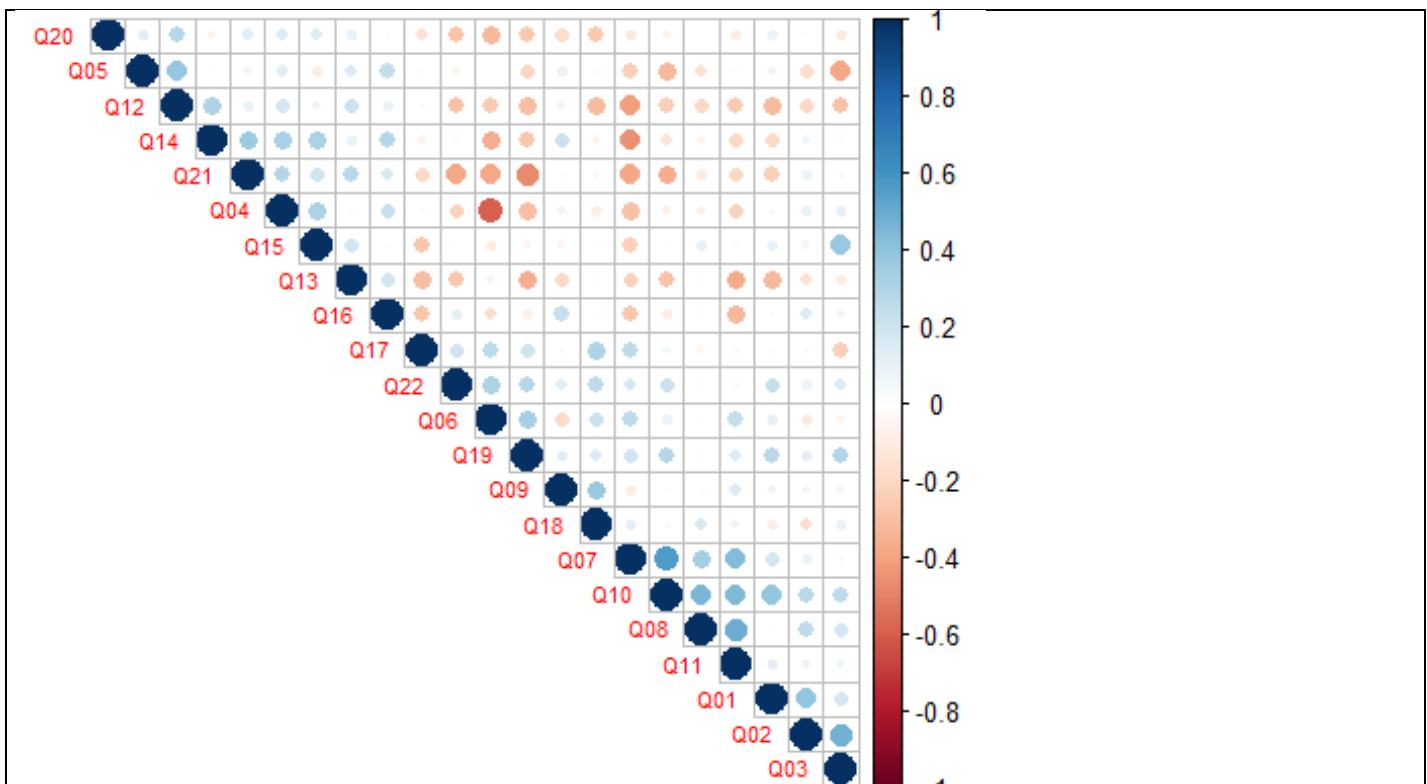
##	Q01	Q02	Q03	Q04	Q05	Q06	Q07	Q08	Q09	Q10	Q11
## Q01	1.00	0.39	0.18	-0.05	-0.07	0.12	0.19	0.00	0.08	0.40	0.13
## Q02	0.39	1.00	0.47	0.09	-0.18	-0.12	0.09	0.25	0.05	0.28	0.08
## Q03	0.18	0.47	1.00	0.10	-0.39	-0.06	-0.04	0.19	0.06	0.26	0.08
## Q04	-0.05	0.09	0.10	1.00	0.14	-0.59	-0.30	-0.08	0.08	-0.08	-0.23
## Q05	-0.07	-0.18	-0.39	0.14	1.00	0.00	-0.25	-0.16	0.09	-0.33	-0.02
## Q06	0.12	-0.12	-0.06	-0.59	0.00	1.00	0.27	0.02	-0.19	0.09	0.25
## Q07	0.19	0.09	-0.04	-0.30	-0.25	0.27	1.00	0.35	-0.11	0.57	0.44
## Q08	0.00	0.25	0.19	-0.08	-0.16	0.02	0.35	1.00	-0.04	0.46	0.50
## Q09	0.08	0.05	0.06	0.08	0.09	-0.19	-0.11	-0.04	1.00	-0.02	0.16
## Q10	0.40	0.28	0.26	-0.08	-0.33	0.09	0.57	0.46	-0.02	1.00	0.45
## Q11	0.13	0.08	0.08	-0.23	-0.02	0.25	0.44	0.50	0.16	0.45	1.00
## Q12	-0.31	-0.20	-0.30	0.19	0.38	-0.26	-0.41	-0.21	0.07	-0.24	-0.25
## Q13	-0.33	-0.17	-0.11	0.03	0.16	-0.07	-0.23	0.02	-0.20	-0.29	-0.37
## Q14	-0.19	0.07	0.03	0.32	0.02	-0.37	-0.45	-0.07	0.21	-0.15	-0.21
## Q15	0.10	0.05	0.38	0.31	-0.10	-0.11	-0.24	0.12	-0.06	0.01	0.03

##	Q16	0.03	0.15	0.07	0.22	0.24	-0.17	-0.27	-0.02	0.24	-0.10	-0.32
##	Q17	-0.03	0.03	-0.25	-0.03	0.02	0.27	0.27	-0.06	0.04	0.06	0.05
##	Q18	-0.11	-0.17	0.09	-0.11	-0.04	0.21	0.12	0.16	0.37	-0.05	0.07
##	Q19	0.28	0.12	0.29	-0.30	-0.21	0.34	0.19	0.02	0.12	0.28	0.15
##	Q20	0.10	-0.01	-0.12	0.16	0.12	-0.33	-0.11	0.00	-0.19	-0.07	-0.10
##	Q21	-0.24	0.09	0.03	0.29	0.07	-0.38	-0.39	-0.11	0.03	-0.37	-0.21
##	Q22	0.23	0.09	0.14	-0.22	-0.06	0.31	0.17	0.02	0.13	0.21	-0.05
##		Q12	Q13	Q14	Q15	Q16	Q17	Q18	Q19	Q20	Q21	Q22
##	Q01	-0.31	-0.33	-0.19	0.10	0.03	-0.03	-0.11	0.28	0.10	-0.24	0.23
##	Q02	-0.20	-0.17	0.07	0.05	0.15	0.03	-0.17	0.12	-0.01	0.09	0.09
##	Q03	-0.30	-0.11	0.03	0.38	0.07	-0.25	0.09	0.29	-0.12	0.03	0.14
##	Q04	0.19	0.03	0.32	0.31	0.22	-0.03	-0.11	-0.30	0.16	0.29	-0.22
##	Q05	0.38	0.16	0.02	-0.10	0.24	0.02	-0.04	-0.21	0.12	0.07	-0.06
##	Q06	-0.26	-0.07	-0.37	-0.11	-0.17	0.27	0.21	0.34	-0.33	-0.38	0.31
##	Q07	-0.41	-0.23	-0.45	-0.24	-0.27	0.27	0.12	0.19	-0.11	-0.39	0.17
##	Q08	-0.21	0.02	-0.07	0.12	-0.02	-0.06	0.16	0.02	0.00	-0.11	0.02
##	Q09	0.07	-0.20	0.21	-0.06	0.24	0.04	0.37	0.12	-0.19	0.03	0.13
##	Q10	-0.24	-0.29	-0.15	0.01	-0.10	0.06	-0.05	0.28	-0.07	-0.37	0.21
##	Q11	-0.25	-0.37	-0.21	0.03	-0.32	0.05	0.07	0.15	-0.10	-0.21	-0.05
##	Q12	1.00	0.22	0.30	0.06	0.09	0.03	-0.31	-0.30	0.28	0.10	-0.29
##	Q13	0.22	1.00	0.09	0.19	0.19	-0.30	-0.03	-0.37	0.09	0.28	-0.27
##	Q14	0.30	0.09	1.00	0.31	0.30	-0.08	-0.07	-0.28	-0.06	0.36	-0.05
##	Q15	0.06	0.19	0.31	1.00	0.03	-0.27	-0.03	-0.06	0.15	0.20	0.00
##	Q16	0.09	0.19	0.30	0.03	1.00	-0.28	-0.02	-0.08	-0.03	0.17	0.12
##	Q17	0.03	-0.30	-0.08	-0.27	-0.28	1.00	0.31	0.21	-0.16	-0.20	0.19
##	Q18	-0.31	-0.03	-0.07	-0.03	-0.02	0.31	1.00	0.14	-0.26	-0.04	0.27
##	Q19	-0.30	-0.37	-0.28	-0.06	-0.08	0.21	0.14	1.00	-0.27	-0.47	0.29
##	Q20	0.28	0.09	-0.06	0.15	-0.03	-0.16	-0.26	-0.27	1.00	0.13	-0.29
##	Q21	0.10	0.28	0.36	0.20	0.17	-0.20	-0.04	-0.47	0.13	1.00	-0.39
##	Q22	-0.29	-0.27	-0.05	0.00	0.12	0.19	0.27	0.29	-0.29	-0.39	1.00

A análise das correlações fica facilitada com uma boa visualização delas.

Matriz de correlações dos itens

```
corrplot(cortcerj, method="circle", order = "hclust", tl.cex=0.7, type = "upper"
)
```



Andy Field et al. (2012) fazem as seguintes recomendações para a análise da matriz de correlações:

- Quaisquer questões que não estejam correlacionadas com nenhuma outra deve ser excluída da análise.
- Verificar se a matriz de correlações tem variáveis que apresentem correlações maiores que 0.3.
- Verificar se existe correlação superior a 0.9, visto que talvez exista problema de multicolinearidade.

Além de realizar a inspeção na matriz de correlações, deve-se realizar o teste de esfericidade de Bartlett e o cálculo do índice de Kaiser-Meyer-Oaklin (KMO). Esses testes dão uma indicação quanto à adequação do uso da análise fatorial. Enquanto o teste de Bartlett tem por objetivo verificar se existe uma estrutura de correlações entre as questões, o cálculo do KMO informa sobre a adequação do tamanho da amostra utilizada.

```
cortest.bartlett(cortcerj, n=nrow(tce.rj.int))
```

```
## $chisq
## [1] 383.3643
##
## $p.value
## [1] 1.150913e-09
##
## $df
## [1] 231
```

O resultado do teste **descarta a possibilidade de que a matriz de correlações seja uma matriz identidade,**

significando isso que existe uma estrutura de correlações subjacente que pode ser explorada pela análise fatorial.

O próximo passo é calcular o índice **KMO**. Recomendam-se valores de KMO no mínimo igual a 0.5, sendo que valores entre 0.5 e 0.7 são considerados medíocres, entre 0.7 e 0.8, bons, entre 0.8 e 0.9 muito bons e superiores a 0.9, ótimos. Hair et al. (2009) sugerem que um valor superior a 0.5 seria aceitável.

KMO(tce.rj.int)

```
## Kaiser-Meyer-Olkin factor adequacy
## Call: KMO(r = tce.rj.int)
## Overall MSA = 0.51
## MSA for each item =
## Q01 Q02 Q03 Q04 Q05 Q06 Q07 Q08 Q09 Q10 Q11 Q12 Q13 Q14 Q15
## 0.45 0.45 0.54 0.45 0.44 0.62 0.72 0.53 0.27 0.61 0.43 0.47 0.55 0.62 0.30
## Q16 Q17 Q18 Q19 Q20 Q21 Q22
## 0.54 0.36 0.36 0.60 0.61 0.69 0.48
```

O valor obtido de **0.51** está no limiar do que se considera adequado de acordo com o autor acima mencionado.

Feitas essas análises preliminares, a extração dos fatores é a próxima etapa. Existem alguns métodos para a realização desta tarefa, sendo aqui utilizado o método das componentes principais sem qualquer rotação. A primeira coisa a fazer é realizar a extração de tantos fatores quanto a quantidade de variáveis (22), o que se faz a seguir:

```
pc1 <- principal(tce.rj.int, nfactors = 22, rotate = "none")
pc1

## Principal Components Analysis
## Call: principal(r = tce.rj.int, nfactors = 22, rotate = "none")
## Standardized loadings (pattern matrix) based upon correlation matrix
##      PC1  PC2  PC3  PC4  PC5  PC6  PC7  PC8  PC9  PC10  PC11
## Q01 0.41 0.37 0.00 0.51 -0.23 0.16 0.01 -0.19 0.27 0.13 -0.20
## Q02 0.21 0.64 0.02 0.23 -0.04 0.00 -0.34 0.09 -0.07 0.45 0.25
## Q03 0.22 0.72 0.16 -0.21 -0.25 -0.14 0.10 -0.10 -0.04 0.02 0.28
## Q04 -0.50 0.41 0.07 0.21 0.26 -0.21 0.01 0.16 0.31 -0.11 0.00
## Q05 -0.36 -0.36 0.09 0.29 0.19 0.52 0.12 -0.09 0.18 0.32 0.06
## Q06 0.56 -0.43 0.08 -0.20 -0.30 0.18 0.20 0.04 -0.01 0.37 -0.04
## Q07 0.71 -0.08 -0.32 -0.03 0.21 0.04 -0.25 0.16 0.11 -0.14 -0.10
## Q08 0.35 0.34 -0.31 -0.37 0.39 0.34 -0.02 0.22 0.02 0.03 0.19
## Q09 0.01 0.14 0.60 0.11 0.50 0.15 0.04 -0.36 -0.10 -0.18 0.06
## Q10 0.63 0.39 -0.25 0.13 0.18 0.13 0.03 0.28 -0.08 -0.19 -0.06
## Q11 0.54 0.10 -0.30 -0.11 0.51 0.17 0.26 -0.27 -0.16 0.20 -0.13
## Q12 -0.60 -0.22 -0.09 0.30 0.18 0.08 0.35 0.24 -0.24 -0.01 0.28
## Q13 -0.50 -0.10 -0.16 -0.48 -0.25 0.31 -0.09 0.18 0.10 -0.03 0.20
## Q14 -0.51 0.32 0.33 -0.05 0.22 -0.09 0.16 0.32 -0.28 0.10 -0.26
## Q15 -0.21 0.53 -0.04 -0.25 -0.13 -0.02 0.63 0.08 0.26 0.10 -0.11
```

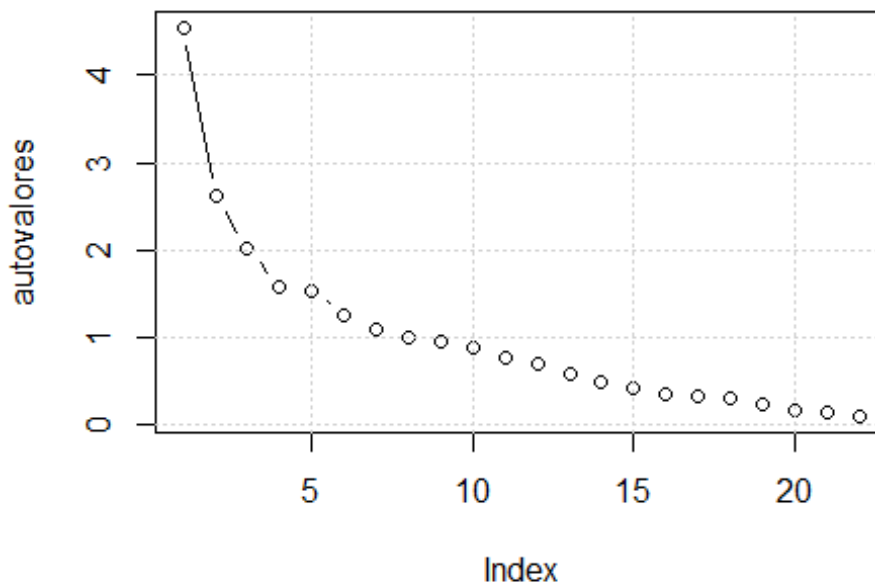
##	Q16	-0.33	0.26	0.43	0.11	-0.10	0.57	-0.27	0.13	-0.03	-0.10	0.04	
##	Q17	0.32	-0.39	0.22	0.21	0.33	-0.41	-0.03	0.37	0.22	0.30	0.20	
##	Q18	0.27	-0.15	0.52	-0.46	0.30	-0.03	-0.01	-0.11	0.44	-0.09	0.14	
##	Q19	0.62	0.04	0.29	0.15	-0.24	-0.10	0.22	-0.13	-0.16	-0.09	0.38	
##	Q20	-0.33	0.09	-0.52	0.32	-0.01	0.04	0.11	-0.14	0.40	-0.14	0.17	
##	Q21	-0.62	0.24	-0.01	-0.28	0.11	-0.17	-0.28	-0.21	0.06	0.34	-0.14	
##	Q22	0.46	0.01	0.51	0.08	-0.20	0.16	0.08	0.33	0.19	-0.08	-0.24	
##		PC12	PC13	PC14	PC15	PC16	PC17	PC18	PC19	PC20	PC21	PC22	h2
##	Q01	0.11	0.31	0.02	-0.17	-0.11	0.08	-0.05	-0.11	0.10	0.03	0.07	1
##	Q02	0.12	-0.02	-0.16	-0.08	0.04	-0.12	0.05	0.02	-0.16	-0.09	0.01	1
##	Q03	-0.09	-0.15	-0.19	0.20	-0.10	0.09	0.16	-0.05	0.15	0.09	-0.01	1
##	Q04	-0.49	0.02	0.00	-0.04	-0.10	-0.07	-0.03	0.13	-0.07	0.09	0.09	1
##	Q05	-0.27	-0.10	-0.09	-0.01	0.16	0.18	0.14	-0.06	-0.01	0.04	-0.07	1
##	Q06	0.01	0.08	0.09	0.14	-0.23	-0.08	0.01	0.11	-0.11	0.16	0.01	1
##	Q07	-0.04	0.16	-0.07	0.22	0.22	-0.15	0.12	-0.18	-0.03	0.10	0.05	1
##	Q08	0.06	-0.18	0.13	-0.24	-0.03	0.01	-0.18	-0.08	0.03	0.12	-0.02	1
##	Q09	0.19	0.19	-0.18	0.02	0.04	-0.14	-0.06	0.10	-0.01	0.11	-0.08	1
##	Q10	-0.01	0.21	0.00	0.15	-0.07	0.30	0.00	0.13	-0.07	-0.06	-0.09	1
##	Q11	-0.09	-0.08	0.00	0.05	0.01	-0.08	0.05	0.13	0.10	-0.13	0.12	1
##	Q12	0.15	0.08	-0.16	0.15	-0.11	0.04	-0.13	-0.14	-0.03	-0.01	0.11	1
##	Q13	0.04	0.35	-0.17	-0.11	0.11	0.00	0.06	0.17	0.08	-0.01	0.06	1
##	Q14	0.21	0.04	0.18	-0.12	0.06	0.05	0.27	-0.01	0.00	0.08	0.04	1
##	Q15	-0.03	0.12	0.02	0.07	0.13	-0.16	-0.08	-0.09	-0.04	-0.08	-0.09	1
##	Q16	-0.07	-0.01	0.31	0.22	-0.08	-0.14	-0.02	-0.02	0.07	-0.09	0.00	1
##	Q17	0.05	0.10	0.08	0.05	0.01	-0.04	-0.02	0.05	0.18	-0.03	-0.06	1
##	Q18	0.12	0.03	0.04	-0.02	-0.13	0.13	0.10	-0.10	-0.11	-0.10	0.06	1
##	Q19	-0.12	0.08	0.29	-0.05	0.28	0.08	-0.05	0.04	-0.04	0.02	0.07	1
##	Q20	0.40	-0.18	0.15	0.10	0.03	-0.03	0.12	0.13	-0.02	0.05	0.01	1
##	Q21	0.09	0.05	0.09	0.24	0.14	0.19	-0.21	0.02	0.00	0.05	0.04	1
##	Q22	0.15	-0.32	-0.22	0.03	0.15	0.04	-0.13	0.10	0.02	0.00	0.07	1
##			u2	com									
##	Q01	-4.4e-16	7.3										
##	Q02	1.0e-15	4.3										
##	Q03	8.9e-16	3.4										
##	Q04	1.1e-16	6.1										
##	Q05	6.7e-16	7.3										
##	Q06	-2.2e-16	5.9										
##	Q07	1.1e-16	3.6										
##	Q08	-6.7e-16	9.6										
##	Q09	5.6e-16	4.6										
##	Q10	1.0e-15	5.0										
##	Q11	-2.2e-16	5.7										
##	Q12	6.7e-16	5.9										
##	Q13	1.7e-15	6.5										
##	Q14	1.4e-15	8.0										
##	Q15	-2.2e-16	4.0										
##	Q16	-2.2e-16	5.7										

```
## Q17 1.3e-15 9.1
## Q18 -1.3e-15 5.8
## Q19 0.0e+00 5.2
## Q20 -4.4e-16 6.6
## Q21 6.7e-16 5.5
## Q22 1.1e-15 6.7
##
##          PC1  PC2  PC3  PC4  PC5  PC6  PC7  PC8  PC9  PC10
## SS loadings  4.54 2.61 2.02 1.57 1.54 1.25 1.09 1.00 0.95 0.87
## Proportion Var 0.21 0.12 0.09 0.07 0.07 0.06 0.05 0.05 0.04 0.04
## Cumulative Var 0.21 0.33 0.42 0.49 0.56 0.62 0.66 0.71 0.75 0.79
## Proportion Explained 0.21 0.12 0.09 0.07 0.07 0.06 0.05 0.05 0.04 0.04
## Cumulative Proportion 0.21 0.33 0.42 0.49 0.56 0.62 0.66 0.71 0.75 0.79
##          PC11 PC12 PC13 PC14 PC15 PC16 PC17 PC18 PC19 PC20
## SS loadings  0.77 0.69 0.58 0.49 0.41 0.35 0.33 0.30 0.23 0.15
## Proportion Var 0.03 0.03 0.03 0.02 0.02 0.02 0.01 0.01 0.01 0.01
## Cumulative Var 0.83 0.86 0.89 0.91 0.93 0.94 0.96 0.97 0.98 0.99
## Proportion Explained 0.03 0.03 0.03 0.02 0.02 0.02 0.01 0.01 0.01 0.01
## Cumulative Proportion 0.83 0.86 0.89 0.91 0.93 0.94 0.96 0.97 0.98 0.99
##          PC21 PC22
## SS loadings  0.14 0.09
## Proportion Var 0.01 0.00
## Cumulative Var 1.00 1.00
## Proportion Explained 0.01 0.00
## Cumulative Proportion 1.00 1.00
##
## Mean item complexity = 6
## Test of the hypothesis that 22 components are sufficient.
##
## The root mean square of the residuals (RMSR) is 0
## with the empirical chi square 0 with prob < NA
##
## Fit based upon off diagonal values = 1
```

A determinação da quantidade de fatores a ser extraída pode dar-se, principalmente, de duas formas: tomando-se todos os fatores cujos autovalores sejam superiores a 1 (critério de Kaiser), ou inspecionando um gráfico denominado screeplot.

```
plot(pc1$values, type = "b", ylab="autovalores", main="Screeplot")
grid()
```

Screepplot



```
#scree(cortcerj)
```

O gráfico evidencia a extração de 7 fatores. Sabendo-se agora a quantidade de fatores a ser extraída, refaz-se a análise explicitando a quantidade de fatores desejada, ainda sem qualquer rotação.

```
pc2 <- principal(tce.rj.int, nfactors = 7, rotate = "none")
pc2
```

```
## Principal Components Analysis
## Call: principal(r = tce.rj.int, nfactors = 7, rotate = "none")
## Standardized loadings (pattern matrix) based upon correlation matrix
##      PC1  PC2  PC3  PC4  PC5  PC6  PC7  h2  u2  com
## Q01  0.41  0.37  0.00  0.51 -0.23  0.16  0.01  0.64  0.36  3.5
## Q02  0.21  0.64  0.02  0.23 -0.04  0.00 -0.34  0.62  0.38  2.1
## Q03  0.22  0.72  0.16 -0.21 -0.25 -0.14  0.10  0.73  0.27  1.9
## Q04 -0.50  0.41  0.07  0.21  0.26 -0.21  0.01  0.58  0.42  3.4
## Q05 -0.36 -0.36  0.09  0.29  0.19  0.52  0.12  0.68  0.32  3.9
## Q06  0.56 -0.43  0.08 -0.20 -0.30  0.18  0.20  0.72  0.28  3.5
## Q07  0.71 -0.08 -0.32 -0.03  0.21  0.04 -0.25  0.72  0.28  1.9
## Q08  0.35  0.34 -0.31 -0.37  0.39  0.34 -0.02  0.74  0.26  5.9
## Q09  0.01  0.14  0.60  0.11  0.50  0.15  0.04  0.66  0.34  2.3
## Q10  0.63  0.39 -0.25  0.13  0.18  0.13  0.03  0.67  0.33  2.5
## Q11  0.54  0.10 -0.30 -0.11  0.51  0.17  0.26  0.76  0.24  3.5
## Q12 -0.60 -0.22 -0.09  0.30  0.18  0.08  0.35  0.66  0.34  2.9
## Q13 -0.50 -0.10 -0.16 -0.48 -0.25  0.31 -0.09  0.69  0.31  3.6
## Q14 -0.51  0.32  0.33 -0.05  0.22 -0.09  0.16  0.56  0.44  3.3
```

```
## Q15 -0.21  0.53 -0.04 -0.25 -0.13 -0.02  0.63 0.81 0.19 2.7
## Q16 -0.33  0.26  0.43  0.11 -0.10  0.57 -0.27 0.78 0.22 3.8
## Q17  0.32 -0.39  0.22  0.21  0.33 -0.41 -0.03 0.62 0.38 5.0
## Q18  0.27 -0.15  0.52 -0.46  0.30 -0.03 -0.01 0.67 0.33 3.4
## Q19  0.62  0.04  0.29  0.15 -0.24 -0.10  0.22 0.60 0.40 2.3
## Q20 -0.33  0.09 -0.52  0.32 -0.01  0.04  0.11 0.51 0.49 2.7
## Q21 -0.62  0.24 -0.01 -0.28  0.11 -0.17 -0.28 0.63 0.37 2.5
## Q22  0.46  0.01  0.51  0.08 -0.20  0.16  0.08 0.56 0.44 2.6
##
##
##          PC1  PC2  PC3  PC4  PC5  PC6  PC7
## SS loadings      4.54 2.61 2.02 1.57 1.54 1.25 1.09
## Proportion Var   0.21 0.12 0.09 0.07 0.07 0.06 0.05
## Cumulative Var   0.21 0.33 0.42 0.49 0.56 0.62 0.66
## Proportion Explained 0.31 0.18 0.14 0.11 0.11 0.09 0.07
## Cumulative Proportion 0.31 0.49 0.63 0.73 0.84 0.93 1.00
##
## Mean item complexity = 3.1
## Test of the hypothesis that 7 components are sufficient.
##
## The root mean square of the residuals (RMSR) is 0.07
## with the empirical chi square 114 with prob < 0.13
##
## Fit based upon off diagonal values = 0.89
```

A próxima etapa é realizar a interpretação dos fatores com vistas a identificar as variáveis latentes subjacente às questões. Para tanto, faz-se necessário realizar a rotação dos fatores, cujo objetivo é a obtenção de um padrão das cargas fatoriais mais claro.

A rotação oblíqua é indicada, uma vez que se espera existir associação entre os fatores.

```
#af_var <- principal(tce.rj.int, nfactors = 7, rotate = "varimax")
#af_var
af_obl <- principal(tce.rj.int, nfactors = 7, rotate = "oblimin")

## Loading required namespace: GPArotation

af_obl

## Principal Components Analysis
## Call: principal(r = tce.rj.int, nfactors = 7, rotate = "oblimin")
## Standardized loadings (pattern matrix) based upon correlation matrix
##      TC1  TC5  TC2  TC4  TC3  TC7  TC6  h2  u2  com
## Q01 -0.22 0.07 -0.14 0.45 -0.43 0.04 0.34 0.64 0.36 3.7
## Q02 0.29 0.14 -0.48 0.25 -0.21 -0.09 0.38 0.62 0.38 4.0
## Q03 0.01 0.05 -0.63 0.08 0.03 0.48 0.13 0.73 0.27 2.0
## Q04 0.67 -0.09 0.03 0.24 -0.02 0.19 0.04 0.58 0.42 1.5
## Q05 -0.09 0.02 0.78 -0.03 -0.03 -0.09 0.34 0.68 0.32 1.4
## Q06 -0.83 0.02 0.03 -0.12 0.12 -0.01 -0.08 0.72 0.28 1.1
```

```

## Q07 -0.15  0.54 -0.20  0.06 -0.06 -0.43 -0.13  0.72  0.28  2.6
## Q08  0.08  0.85 -0.08 -0.24  0.09  0.06  0.09  0.74  0.26  1.3
## Q09  0.19  0.14  0.21  0.39  0.55  0.04  0.34  0.66  0.34  3.5
## Q10 -0.07  0.61 -0.18  0.28 -0.20  0.01  0.07  0.67  0.33  1.9
## Q11 -0.07  0.82  0.16  0.14  0.08  0.09 -0.18  0.76  0.24  1.3
## Q12  0.20 -0.14  0.69  0.08 -0.15  0.20 -0.09  0.66  0.34  1.6
## Q13 -0.03 -0.08  0.07 -0.79 -0.01  0.09  0.14  0.69  0.31  1.1
## Q14  0.44 -0.14  0.09  0.09  0.28  0.37  0.12  0.56  0.44  3.3
## Q15 -0.01  0.09 -0.03 -0.06 -0.04  0.90 -0.10  0.81  0.19  1.1
## Q16  0.06 -0.10  0.10 -0.15  0.08 -0.04  0.85  0.78  0.22  1.1
## Q17  0.02 -0.10  0.06  0.50  0.31 -0.33 -0.37  0.62  0.38  3.6
## Q18 -0.12  0.12 -0.11 -0.03  0.79 -0.02 -0.02  0.67  0.33  1.1
## Q19 -0.53 -0.09 -0.22  0.42  0.06  0.14  0.00  0.60  0.40  2.6
## Q20  0.25  0.08  0.28 -0.01 -0.58  0.08 -0.10  0.51  0.49  2.0
## Q21  0.64 -0.14 -0.19 -0.34  0.14  0.01  0.01  0.63  0.37  2.0
## Q22 -0.52 -0.11 -0.12  0.28  0.23  0.06  0.32  0.56  0.44  3.1
##
##
##          TC1  TC5  TC2  TC4  TC3  TC7  TC6
## SS loadings      2.79 2.39 2.21 2.06 1.84 1.71 1.64
## Proportion Var   0.13 0.11 0.10 0.09 0.08 0.08 0.07
## Cumulative Var   0.13 0.24 0.34 0.43 0.51 0.59 0.66
## Proportion Explained 0.19 0.16 0.15 0.14 0.13 0.12 0.11
## Cumulative Proportion 0.19 0.35 0.51 0.65 0.77 0.89 1.00
##
## With component correlations of
##          TC1  TC5  TC2  TC4  TC3  TC7  TC6
## TC1  1.00 -0.17  0.15 -0.14 -0.05  0.21  0.11
## TC5 -0.17  1.00 -0.22  0.15 -0.03 -0.04 -0.05
## TC2  0.15 -0.22  1.00 -0.14  0.05 -0.06 -0.06
## TC4 -0.14  0.15 -0.14  1.00  0.03 -0.05  0.07
## TC3 -0.05 -0.03  0.05  0.03  1.00 -0.04  0.01
## TC7  0.21 -0.04 -0.06 -0.05 -0.04  1.00  0.15
## TC6  0.11 -0.05 -0.06  0.07  0.01  0.15  1.00
##
## Mean item complexity = 2.1
## Test of the hypothesis that 7 components are sufficient.
##
## The root mean square of the residuals (RMSR) is 0.07
## with the empirical chi square 114 with prob < 0.13
##
## Fit based upon off diagonal values = 0.89

```

Objetivando facilitar a interpretação dos fatores, será apresentado o resultado acima omitindo-se as cargas fatoriais inferiores a 0.5. Assim, serão mostradas apenas as contribuições (superiores a 0.5) que cada variável faz para cada fator.

```
#print.psych(af_var, cut=0.5, sort = TRUE)
```



```

print.psych(af_obl, cut=0.5, sort = TRUE)

## Principal Components Analysis
## Call: principal(r = tce.rj.int, nfactors = 7, rotate = "oblimin")
## Standardized loadings (pattern matrix) based upon correlation matrix
##      item  TC1  TC5  TC2  TC4  TC3  TC7  TC6  h2  u2  com
## Q06    6 -0.83
## Q04    4  0.67
## Q21   21  0.64
## Q19   19 -0.53
## Q22   22 -0.52
## Q14   14
## Q08    8      0.85
## Q11   11      0.82
## Q10   10      0.61
## Q07    7      0.54
## Q05    5      0.78
## Q12   12      0.69
## Q03    3     -0.63
## Q02    2
## Q13   13      -0.79
## Q17   17
## Q01    1
## Q18   18      0.79
## Q20   20     -0.58
## Q09    9      0.55
## Q15   15      0.90
## Q16   16      0.85 0.78 0.22 1.1
##
##              TC1  TC5  TC2  TC4  TC3  TC7  TC6
## SS loadings      2.79 2.39 2.21 2.06 1.84 1.71 1.64
## Proportion Var   0.13 0.11 0.10 0.09 0.08 0.08 0.07
## Cumulative Var   0.13 0.24 0.34 0.43 0.51 0.59 0.66
## Proportion Explained 0.19 0.16 0.15 0.14 0.13 0.12 0.11
## Cumulative Proportion 0.19 0.35 0.51 0.65 0.77 0.89 1.00
##
## With component correlations of
##      TC1  TC5  TC2  TC4  TC3  TC7  TC6
## TC1  1.00 -0.17  0.15 -0.14 -0.05  0.21  0.11
## TC5 -0.17  1.00 -0.22  0.15 -0.03 -0.04 -0.05
## TC2  0.15 -0.22  1.00 -0.14  0.05 -0.06 -0.06
## TC4 -0.14  0.15 -0.14  1.00  0.03 -0.05  0.07
## TC3 -0.05 -0.03  0.05  0.03  1.00 -0.04  0.01
## TC7  0.21 -0.04 -0.06 -0.05 -0.04  1.00  0.15
## TC6  0.11 -0.05 -0.06  0.07  0.01  0.15  1.00
##
## Mean item complexity = 2.1

```

```
## Test of the hypothesis that 7 components are sufficient.  
##  
## The root mean square of the residuals (RMSR) is  0.07  
## with the empirical chi square 114 with prob <  0.13  
##  
## Fit based upon off diagonal values = 0.89
```

Tomando por base o resultado da análise empreendida obtido com a rotação oblíqua, procede-se a seguir à identificação dos fatores subjacentes:

Fator 1:

Q06 - O Tribunal utiliza *dashboard* para monitorar métricas-chaves para o controle.

Q04 - As iniciativas de uso de análise de dados são isoladas.

Q21 - O uso de mineração de texto para auxiliar no exercício do controle externo ainda não é uma realidade.

Q19 - O uso de softwares que suportam análise de dados é difundido entre os auditores.

Q22 - O Tribunal dispõe de técnicos capacitados em softwares analíticos.

Examinando estes itens, pode-se sugerir que o fator se relaciona à **existência de softwares e técnicas analíticas em conjunto com pessoas capacitadas**.

Fator 2:

Q08 - O Tribunal possui de um setor que tem como missão gerenciar informações para dar suporte às atividades de controle externo do Tribunal.

Q11 - O Tribunal possui em sua estrutura organizacional um setor encarregado de garantir a qualidade dos dados utilizados para a realização das auditorias/fiscalizações.

Q10 - A análise de dados é utilizada para auxiliar na detecção de fraudes.

Q07 - Além dos dados obtidos com os jurisdicionados, o Tribunal também disponibiliza aos técnicos dados obtidos de outras instituições.

Este fator parece estar associado à **existência e uso de dados**.

Fator 3:

Q05 - Inexiste incentivo por parte do Tribunal para que os técnicos se capacitem em análise de dados.

Q12 - Inexiste apoio da alta administração do Tribunal ao uso de análise de dados para realizar a atividade de controle.

Q03 - O Tribunal oferece a seus técnicos capacitação em análise de dados.

Este fator se relaciona a **apoio institucional para o uso de análise de dados**.

Fator 4:

Q13 - Os recursos computacionais (hardware) disponibilizados aos técnicos são inadequados para trabalhar com grandes bases de dados

Este fator se relaciona a **trabalhar com grandes bases de dados**

Fator 5:

Q18 - O Tribunal possui ações estratégicas previstas explicitamente em seu Planejamento Estratégico visando aumentar sua capacidade analítica.

Q20 - Os técnicos têm dificuldade de dispor dos dados necessários à realização de auditorias/fiscalizações.

Q09 - A análise de risco é feita de forma subjetiva, valendo-se da experiência do técnico.

Este fator parece não ter uma interpretação muito clara, mas talvez seja possível vinculá-lo à **falta de vontade institucional para o uso de análise de dados**.

Fator 6:

Q15 - Os técnicos não se sentem confortáveis para trabalhar com grandes bases de dados.

Este fator pode ser vinculado à **competência técnica para manuseio de grandes bases de dados**.

Fator 7

Q16 - A seleção de objetos de fiscalização é realizada com base na experiência dos técnicos, sem o auxílio de análise de risco.

Este fator poderia indicar a **tomada de decisões sem base em dados**.

Com isso, seriam removidas da escala final as seguintes questões: **Q01, Q02, Q14, Q17**.

Em resumo, tem-se a seguinte composição dos fatores:

Fator	Itens
1	Q04, Q06, Q19, Q21, Q22
2	Q07, Q08, Q10, Q11
3	Q03, Q05, Q12
4	Q13
5	Q09, Q18, Q20
6	Q15
7	Q16

Cálculo da confiabilidade.

A confiabilidade de um instrumento de medida diz respeito à sua habilidade de medir de forma consistente.

A medida estatística mais comumente utilizada para medir a confiabilidade do instrumento de pesquisas é o **alpha de Cronbach**. Esta medida informa em que grau os itens do questionário estão inter-relacionados. O coeficiente varia de 0 a 1 e fornece uma medida da consistência interna de um teste ou escala. A consistência interna de um instrumento descreve em que medida os itens do instrumento estão medindo o

mesmo conceito ou constructo.

Valores dessa estatística superiores a 0.7 são considerados aceitáveis. No questionário-piloto aplicado a técnicos do Tribunal de Contas do Estado do Rio de Janeiro, obteve-se um α de 0.77, o que nos permite concluir pela consistência dos itens.

```
confiab <- alpha(tce.rj.int,  
                keys = c("Q04", "Q05", "Q09", "Q12", "Q13", "Q14", "Q15", "Q16",  
                        "Q20", "Q21"))  
  
confiab$total  
  
## raw_alpha std.alpha G6(smc) average_r S/N ase mean  
## 0.7796779 0.7846707 0.8947888 0.1421012 3.644049 0.04482939 2.460111  
## sd  
## 0.3864341
```

Aplicação da escala nos demais Tribunais de Contas

Construída a escala, o passo seguinte é sua aplicação aos técnicos dos demais Tribunais de Contas com vistas a derivar a pontuação dessas instituições na escala.

A execução desta etapa da pesquisa restou prejudicada em razão da baixa adesão obtida dos Tribunais de Contas, com retorno de poucas respostas, fato que prejudica o cálculo da pontuação dos Tribunais de Contas, já que esta tomaria por base a percepção de poucos técnicos do Tribunal de Contas.

Não obstante, objetivando ilustrar o procedimento de cálculo, será utilizada escala obtida em um Tribunal de Contas com vistas a obter o seu escore. O objetivo é quantificar, com base na escala construída, o escore para representar o resultado da aplicação da escala no referido Tribunal de Contas.

A quantidade de respondentes neste TC é de 5 pessoas. O escore do TC será a média dos escores obtidos por cada um dos técnicos.

```
# Escores dos técnicos  
pts_ind_tc1 <- apply(tce.pe.int, 1, sum)  
# Escores do TC  
pts_tc1 <- mean(pts_ind_tc1)  
pts_tc1  
  
## [1] 56.8
```

O referido TC teve um escore de 56.8. Este escore é alto ou baixo? A escala possui 18 itens, então cada respondente pode ir de um mínimo de 18 pontos até um máximo de 90 pontos. O valor neutro da escala implica uma pontuação de 54. O escore do TC em questão indica um nível de maturidade satisfatório.

Naturalmente que quanto mais técnicos responderem à escala mais confiável será o escore obtido. A mesma análise será feita para um outro TC que apresentou 19 respostas.

```
# Escores dos técnicos
pts_ind_tc4 <- apply(tcu, 1, sum)
# Escores do TC
pts_tc4 <- mean(pts_ind_tc4)
pts_tc4

## [1] 58.47368
```

Este outro TC apresentou um escore maior, no valor de 58,47.

Mas por que o 2º TC teve um escore maior que o primeiro? O que o primeiro TC deve fazer para melhorar?

Calculando-se o escore de cada um dos TCs em cada fator, pode dar uma indicação de onde o TC deve melhorar.

```
calcula_escore <- function(x, escore.rel=FALSE){
  #x - data frame contendo os dados do TC
  escore_fator1 <- mean(apply(x[c("Q06", "Q04", "Q21", "Q19", "Q22")], 1, sum))
  escore_fator2 <- mean(apply(x[c("Q08", "Q11", "Q10", "Q07")], 1, sum))
  escore_fator3 <- mean(apply(x[c("Q05", "Q12", "Q03")], 1, sum))
  escore_fator4 <- mean(apply(x["Q13"], 1, sum))
  escore_fator5 <- mean(apply(x[c("Q18", "Q20", "Q09")], 1, sum))
  escore_fator6 <- mean(apply(x["Q15"], 1, sum))
  escore_fator7 <- mean(apply(x["Q16"], 1, sum))
  escore_total <- mean(apply(x, 1, sum))

  scores <- c(F1=escore_fator1, F2=escore_fator2, F3=escore_fator3,
             F4=escore_fator4, F5=escore_fator5, F6=escore_fator6,
             F7=escore_fator7, Total=escore_total)

  scores_percent <- scores / scores['Total']
  # Apresentar os escores máximos total e por fator...
  if(escore.rel) return(scores_percent) else return(scores)
}
```

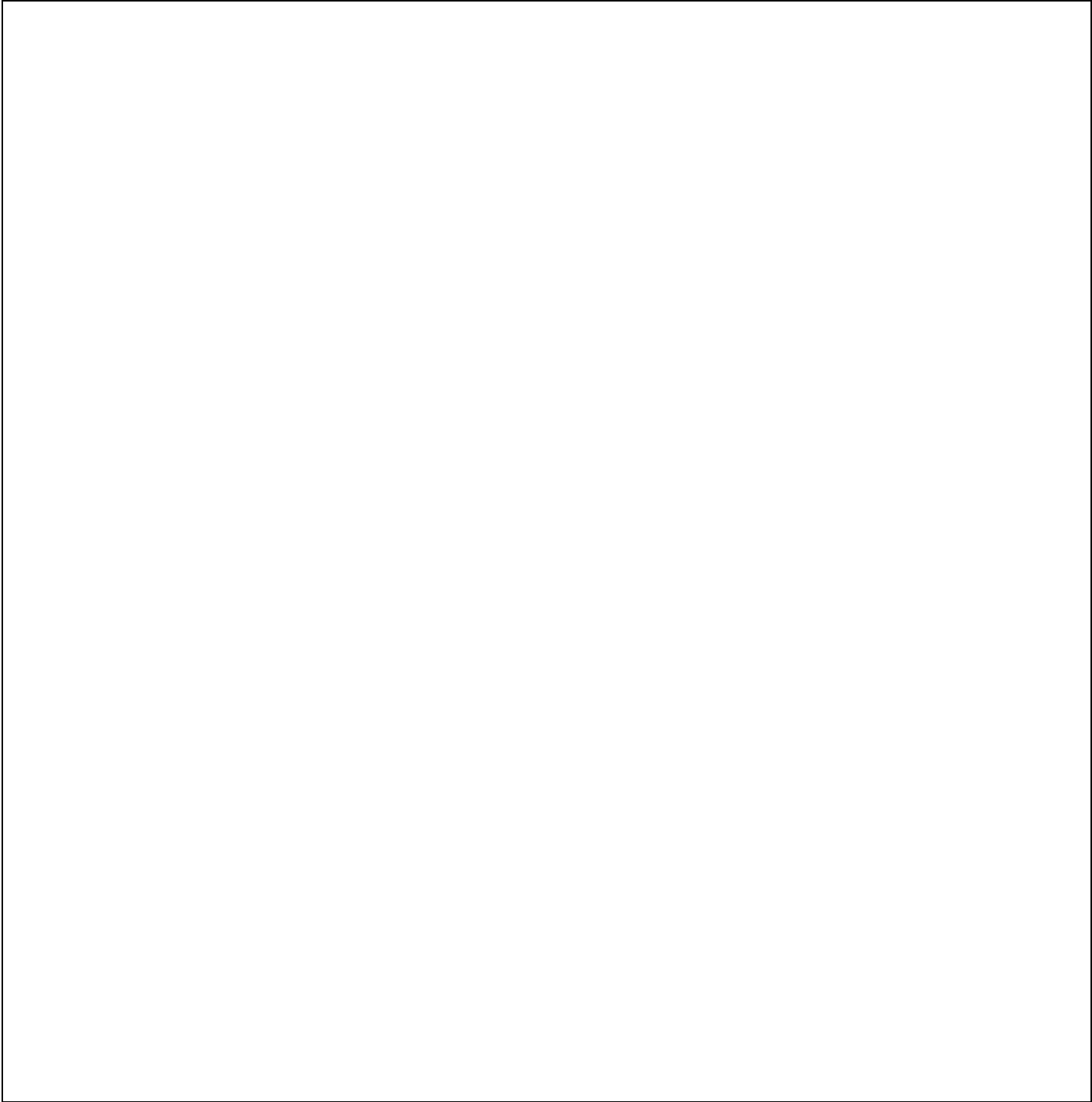
```
round(calcula_escore(tce.pe.int), 2)
```

```
##   F1   F2   F3   F4   F5   F6   F7 Total
## 16.2 15.6  9.2  2.6  7.8  3.2  2.2 56.8
```

```
round(calcula_escore(tcu), 2)
```

```
##   F1   F2   F3   F4   F5   F6   F7 Total
## 16.00 15.32  8.47  2.53 10.00  3.42  2.74 58.47
```

O segundo TC teve escores maiores nos fatores 5, 6 e 7. O primeiro TC, por sua vez, teve escores um pouco maiores nos fatores 1, 2, 3 e 4.



6 - PARECER DO PESQUISADOR-AVALIADOR

Cumprimento do estabelecido no projeto apresentado

Em sua proposta, o projeto de pesquisa "Controle externo baseado em análise de dados: Estão os tribunais de contas preparados?" apresentou o objetivo geral de mensurar o nível de maturidade analítica dos tribunais de contas brasileiros e a sua adequação à adoção de uma estratégia de controle baseada em análise de dados, por meio do atingimento dos objetivos específicos de:

- Criação de uma escala de maturidade aplicável aos tribunais de contas;
- Mensuração do nível de maturidade dos TCs a partir da escala criada;
- Definição dos contornos e requisitos de uma estratégia de controle baseada em análise de dados;
- Avaliação da adequação do nível de maturidade analítica observada nos TCs com uma estratégia de controle calcada em análise de dados;
- Identificação de oportunidades de melhorias no nível de maturidade analítica nos TCs pesquisados e, notadamente, no TCE-RJ.

Tais objetivos foram plenamente atingidos conforme será visto adiante.

Referencial teórico

Baseada em um referencial teórico abrangente, para o atingimento de seus objetivos a pesquisa buscou respaldo em publicações que abordam o uso de dados como diferencial estratégico (competição analítica, mineração de dados, competência analítica etc.) tanto na esfera pública quanto no meio privado, sendo este último o ambiente onde tais técnicas se encontram em estágio mais avançado.

A presença de referências da pesquisa a trabalhos do Tribunal de Contas da União denota a importância que tal Corte de Contas vem dando ao tema da utilização de dados no suporte às atividades de fiscalização e controle.

Há ainda referências que suportam as etapas da pesquisa e as técnicas empregadas no ajuste do instrumento de coleta e na análise feita nos dados obtidos.

Dada a natureza da pesquisa, a qual permeia conceitos de análises de dados e de controle externo, tais referências bibliográficas estão alinhadas com o objetivo do trabalho.

Enfoque Metodológico

A partir do levantamento das referências bibliográficas, os pesquisadores optaram por utilizar uma escala de Likert (escala psicométrica comumente usada em pesquisas de opinião) para a mensuração do conceito de maturidade analítica, partindo de itens vistos no referencial teórico, os quais foram validados com especialistas.

As etapas da pesquisa abrangeram a (1) Definição do Conceito a ser medido pela pesquisa, que é a maturidade analítica de uma organização; (2) Geração inicial dos itens constantes do questionário; (3) Validação com especialistas; (4) Aplicação preliminar do questionário para verificar a necessidade de ajustes; (5) Aplicação final do questionário e, finalmente, a (6) Análise dos Dados.

Na aplicação do questionário junto a representantes de outros tribunais de contas, a equipe constatou que, mesmo com o caráter anônimo da pesquisa em não identificar os respondentes e nem os tribunais, pode ter havido um certo desconforto de alguns auditores para responder à pesquisa. Tal hipótese explicaria a baixa adesão dos integrantes de alguns tribunais em preencher o questionário.

O atingimento dos objetivos da pesquisa denota a adequação do procedimento metodológico empregado no presente trabalho.

Contribuições e Resultados da Pesquisa

A pesquisa criou, de forma estruturada, um instrumento capaz de medir os tribunais de contas quanto à sua maturidade analítica, ou seja, a sua capacidade de nortear suas ações de controle a partir de análises de dados.

Tal instrumento se desdobra em 7 fatores (“Existência de pessoal capacitado em softwares e técnicas analíticas”, “Existência e uso de dados”, “Apoio institucional para o uso de análise de dados”, “Infraestrutura adequada para manipular grandes bases de dados”, “Vontade institucional para o uso de análise de dados”, “Competência técnica para manuseio de grandes bases de dados” e “Tomada de decisões baseada em dados (fatos)”.

A "régua" fruto da pesquisa pode contribuir significativamente no diagnóstico sobre como se encontra um tribunal de contas no que diz respeito a uma capacidade cada vez mais estratégica, em um momento no qual se exige gradativamente mais efetividade dos órgãos de controle.

Além de prover uma fotografia sobre o grau de maturidade analítica de uma organização, é possível se extrair do emprego da pesquisa um direcionamento sobre como avançar no aprimoramento da atividade de controle externo baseada em análise de dados.

A partir dos dados obtidos pela pesquisa, verificou-se a maturidade analítica de 4 dos Tribunais de Contas, os quais apresentaram 5 ou mais respondentes.

A abertura dos resultados da pesquisa por fator (pessoal especializado, apoio institucional, existência de dados etc.) permite a um Tribunal respondente verificar as áreas nas quais ele possui mais oportunidade de melhoria.

Pesquisador

Nome: Marcos Ferreira da Silva

Rio de Janeiro, 31 de agosto de 2017.

Pesquisador-coordenador

Nome: Bruno Mattos Souza de Souza Melo

Rio de Janeiro, 31 de agosto de 2017.

(assinatura do Pesquisador-coordenador)

Pesquisador-avaliador

Nome: Alberto de Fontes Tavares Neto

Rio de Janeiro, 31 de agosto de 2017.
