**Universidade de São Paulo**

**Faculdade de Filosofia, Letras e Ciências Humanas**

**Departamento de Ciência Política**

**FLS 5028 – Métodos Quantitativos e Técnicas de Pesquisa em Ciência Política**

**FLP 0406 – Métodos e Técnicas de Pesquisa em Ciência Política**

**1º Semestre de 2019**

**Profº. Dr. Glauco Peres da Silva**

**LISTA DE EXERCÍCIOS 03**

**"Desenho de Pesquisa Empírica em Ciência Política"**

**Data de entrega: 25/03/2019 (noturno) e 27/03/2019 (vespertino)**

**Nome: Kaue Oliveira Almeida | NUSP: 10763438**

**Período: (X) Vespertino, () Noturno**

**Exercícios Regulares**

**Exercício 01**

Com base na bibliografia obrigatória indicada, marque “Verdadeiro” (V) ou “Falso” (F) para as afirmações a seguir e, quando falsas, justifique sua escolha em 5 linhas, no máximo.

(V) Um passo importante em estudos empíricos é a forma como saímos do nível teórico para a mensuração. Esse é um aspecto central seja na coleta de dados pelo pesquisador ou na utilização de dados disponíveis em diversas fontes.

Justificativa: É fundamental para o atestamento de uma teoria que a mensuração do estudo esteja de acordo com os conceitos estabelecidos na teoria. Sobretudo nas ciências sociais, muitos conceitos são difíceis de serem observados de forma empírica e diretamente. Dessa forma, é essencial para o estudo que os dados a serem utilizados sejam avaliados anteriormente a teste de hipótese.

(F) Dois pontos de atenção na operacionalização de uma variável são sua confiabilidade e o viés de mensuração. A primeira indica se, com a aplicação dos mesmos procedimentos, se mensurarmos os mesmos casos ou observações, os resultados produzidos seriam idênticos. Já o viés de mensuração indica se existe subestimação ou sobre-estimação de modo sistemático nos valores de uma variável. Se buscamos testar teorias, devemos priorizar a redução do viés à confiabilidade da mensuração.

Justificativa: A verificação da confiabilidade e do viés de mensuração são sim passos para mensuração, mas o viés de mensuração não é decisivo para o acolhimento da medida. Apesar de o viés ser um problema relevante na mensuração, quando ela é sobre-estimada, o mesmo padrão será visível tanto para variável independente quanto para a variável dependente. Mas se a mensuração de qualquer variável não é confiável, a relação entre as variáveis se torna obscura.

(V) No processo de observar se uma medida representa teoricamente o que buscamos analisar é necessário observar sua validade, ou seja, se ela mede o que foi pretendido inicialmente. Nesse sentido, ao observamos que uma medida é confiável devemos avaliar sua validade em três aspectos: face, conteúdo e constructo. A partir disso, com medidas válidas e confiáveis podemos realizar nossas avaliações sobre as teorias causais.

Justificativa: Um passo importante para a mensuração, assim como o atestamento da confiabilidade, é assegurar que a medida realmente corresponde ao conceito de nossa teoria. Apesar de todos os meios de avaliar a validade de uma medida serem teóricos e sujeitos a graus de contestação, dependemos de uma série de meios sobrepostos para atestá-la. Estes meios são a verificação: da validade pela face, da validade de conteúdo e a validade de constructo.

(F) A capacidade de capturar no mundo real o que espelhamos em conceitos abstratos não interfere diretamente na capacidade de avaliar o suporte empírico para uma teoria.

Justificativa: Se não podemos capturar no mundo real o que espelhamos em conceitos abstratos, nossa capacidade de avaliar de modo conclusivo o suporte empírico fica comprometida. Pois se não podemos determinar com confiança que nossas medidas correspondem aos conceitos de nossa teoria, da mesma forma não podemos determinar com confiança que as afirmações acerca da teoria pelo estudo são coerentes.

(V) Para distribuições assimétricas, é recomendável utilizar a mediana ao invés da média como medida de tendência central. Uma distribuição com obliquidade (*skewness*) positiva possui mais valores acima da média do que abaixo. Já uma distribuição com obliquidade negativa possui mais valores abaixo da média do que acima. A diferença entre a média e a mediana é uma das maneiras de se analisar a obliquidade de uma variável.

Justificativa: Em distribuições assimétricas a média é afetada pelos valores extremos, assim é mediana é mais recomendável. Uma das formas que podemos verificar a obliquidade de uma variável é calcular a diferença entre a média e a mediana dos valores.

(F) Variáveis categóricas são variáveis que não é possível de se fazer uma ordenação que distinga cada um de seus valores possíveis. Assim sendo, somente podemos usar a média e a moda para descrever suas distribuições. Já variáveis contínuas são variáveis que possuem uma diferença de unidade constante, podendo ser utilizadas tanto a média, a moda e a mediana para descrevê-las.

Justificativa: A única medida apropriada para a mensuração da tendência central de uma variável categórica é a moda.

(V) Apesar de semelhantes, gráficos de barras e histogramas são adequados para diferentes tipos de variáveis. Os gráficos de barras são utilizados para descrever variáveis categóricas e os histogramas para variáveis contínuas.

Justificativa: O gráfico de barras descreve a frequência de variáveis categóricas. Já os histogramas descrevem a distribuição de variáveis contínuas.

(F) A utilização das estatísticas e dados descritivos nos permitem realizar inferências causais sobre as relações teóricas que buscamos observar. Ao conhecer os dados, conseguimos identificar os padrões de relação entre as variáveis e identificar se uma teoria possui respaldo empírico.

Justificativa: Os dados descritivos nos fornecem informações de apenas uma variável. Assim, eles são importantes para conhecermos de uma forma melhor nossos dados, mas não são suficientes para testarmos nossas teorias. Isso se deve pelo fato de que uma teoria é uma proposição acerca de uma relação causal entre duas variáveis.

**Exercício 02**

Um conceito fundamental para a ciência política é o de democracia. Como apontado por Kellstedt e Whitten (2015: 124-129), todavia, esse é um conceito que é difícil de se operacionalizar no mundo real. Uma medida de mensuração bastante difundida por pesquisadores da área é o *Polity IV*. Os pesquisadores ligados a esse projeto criaram uma escala que vai de -10 a +10, atribuindo um *score* a cada país, operacionalizando 4 componentes de seus governos: i) regras para o recrutamento do Executivo; ii) competitividade do recrutamento do Executivo; iii) abertura do recrutamento do Executivo; e iv) limites legais ao chefe do Executivo.

Por outro lado, temos também medidas menos complexas que decorrem de uma definição minimalista de democracia, ou é uma democracia ou não é. Nesse sentido, outras medidas também são operacionalizadas como a de Cheibub, Gandhi e Vreeland (2010), o Democracy-Dictatorship (DD) Index assume a dicotomia em apenas dois valores possíveis (1 para democracia, 0 para ditadura).

(A) As duas medidas lidam com o mesmo conceito (democracia) a partir de diferentes concepções teóricas (poliarquia e definição minimalista de democracia), o que tem implicações sobre como as medidas são obtidas. Que tipo de variável (categórica, ordinal ou contínua) é cada uma dessas variáveis usadas para mensurar a democracia?

Dica: pense no significado dos zeros.

(Utilize no máximo 10 linhas para a realização dessa questão)

A) A variável *Polity IV* é do tipo contínua, pois ela é uma escala em que há uma diferença de unidade constante, ou seja, se aumentarmos o *Polity IV* em +1 de um valor +2, isso significará a mesma coisa se aumentarmos também em +1 em um valor +3 de *Polity IV*.

A variável DD Index é do tipo categórica, pois os valores que podem ser assumidos, 0 (ditadura) e 1 (democracia), são diferentes e podemos categorizá-los, mas não é possível fazer qualquer ordem ou ranqueamento entre eles.

(B) De acordo com sua resposta no item (A), justifique que tipo de medida de tendência central você usaria para analisar o *Polity IV* e o *DD Index*. (Utilize no máximo 5 linhas para a realização dessa questão)

B) A única medida de tendência central apropriada para o DD Index, tendo em vista que é uma variável categórica, é a moda. Já para a variável *Polity IV*, a mediana é uma medida de tendência central mais confiável do que a média, outra medida para variáveis contínuas, pois os valores extremos não interferem em seu cálculo.

(C) De acordo com sua resposta no item (A), justifique que tipo de gráfico você usaria para descrever os dados do *Polity IV* e o *DD Index*. (Utilize no máximo 5 linhas para a realização dessa questão)

C) O gráfico mais apropriado para o DD Index é o gráfico de barras, tendo em vista que é uma variável categórica e o uso de gráfico de setores não é recomendado por muitos estatísticos. Para a variável *Polity IV,* o gráfico a ser utilizado seria o box-plot, onde seria possível verificar a tendência central, pela mediana, bem como a distribuição e os valores discrepantes da variável.

**Exercícios de fechamento de Bloco**

**Tópicos da Aula 1**

A partir da leitura do seguinte excerto retirado da introdução da Democracia na América de Tocqueville responda às seguintes perguntas:

Entre os novos objetos que me chamaram a atenção durante minha permanência nos Estados Unidos nenhum me impressionou mais do que a igualdade das condições. Descobri sem custo a influência prodigiosa que exerce esse primeiro fato sobre o andamento da sociedade; ele proporciona ao espírito público certa direção, certo aspecto às leis; aos governantes, novas máximas e hábitos particulares aos governados.

Não tardei a reconhecer que esse mesmo fato estende sua influência muito além dos costumes políticos e das leis, e tem império sobre a sociedade tanto quanto sobre o governo: cria opiniões, faz nascer sentimentos, sugere usos e modifica tudo o que ele não produz. Assim, pois, à medida que eu estudava a sociedade americana, via cada vez mais, na igualdade das condições, o fato gerador de que cada fato particular parecia decorrer e deparava incessantemente com ele como um ponto central a que todas as minhas observações confluíam. Dirigi então meu pensamento para nosso hemisfério e pareceu-me que distinguia nele algo análogo ao espetáculo que o novo mundo me oferecia. (...).

Uma grande revolução democrática se realiza entre nós; todos a vêem, mas nem todos a julgam da mesma maneira. (...). Transporto-me por um momento ao que era a França há setecentos anos: encontro-a dividida entre um pequeno número de famílias que possuem a terra e governam os habitantes; o direito de comandar descende então de geração em geração com as heranças; os homens têm um único meio de agir uns sobre os outros, a força; descubro uma só origem para o poder, a-propriedade da terra. Mas eis que o poder político do clero vem se estabelecer e, logo, se ampliar. O clero abre-se a todos, ao pobre e ao rico, ao plebeu e ao senhor; a igualdade começa a penetrar pela Igreja no seio do governo, e aquele que vegetará como servo numa eterna escravidão coloca-se como padre no meio dos nobres e muita vez vai sentar-se acima dos reis. Tomando-se a sociedade, com o passar do tempo, mais civilizada e mais estável, as diferentes relações entre os homens se tornam mais complicadas e mais numerosas. A necessidade das leis civis faz-se sentir vivamente. Nascem então os legistas; eles saem do recinto obscuro dos tribunais e do reduto poeirento dos cartórios e vão ocupar um lugar na corte do príncipe, ao lado dos barões feudais cobertos de arminho e de ferro. Arruínam-se os reis nos grandes empreendimentos; esgotam-se os nobres nas guerras privadas; enriquecem-se os plebeus no comércio. A influência do dinheiro começa a fazer-se sentir sobre os negócios do Estado. O negócio é uma nova fonte que se abre ao poder, e os financistas se tornam um poder político que os demais menosprezam e adulam. Pouco a pouco, as luzes se disseminam; vê-se despertar o gosto pela literatura e pelas artes; o espírito se torna então um elemento de sucesso; a ciência é um meio de governo, a inteligência uma força social; os letrados chegam aos negócios. Entretanto, à medida que se descobrem novos caminhos para chegar ao poder, vê-se diminuir o valor do nascimento. No século XI, a nobreza tinha um preço inestimável; é comprada no século XIII; o primeiro enobrecimento ocorre em 1270, e a igualdade se introduz enfim no governo pela própria aristocracia. (TOCQUEVILLE, Alexis de. A Democracia na América, São Paulo: Martins Fontes, 2005.)

(A) Tocqueville está tratando do surgimento de uma era democrática tanto no que diz respeito da forma de governo e também na forma como a sociedade se organiza. Com base no excerto apresentado acima apresente o tipo de causalidade que o autor utiliza, justifique sua resposta em no máximo 10 linhas.

A) O tipo de causalidade apresentado pelo autor é do tipo contrafactual. Isto se explica pelo fato de ao tentar demonstrar que a igualdade das condições na América está inaugurando uma era democrática, para sua argumentação, ele faz o uso de um cenário sem a igualdade das condições. Assim, ele toma a igualdade das condições como causa e o surgimento da democracia participativa como efeito. Desse modo, Tocqueville verifica que a igualdade de condições está ocorrendo na América, então há o surgimento da democracia participativa. Contrafactualmente, verifica historicamente que pela ausência de igualdade de condições na França não houve democracia participativa. Porém, com a abertura do poder político para outras classes, houve, gradualmente, o surgimento da democracia participativa.

(B) Apresente um desenho de pesquisa que busque explicar o mesmo fenômeno, ou seja o surgimento da democracia. Crie uma teoria que contenha todos os elementos que foram explicitados na capítulo Panorama geral nas Ciências Sociais. Ou seja, a sua explicação deve conter elementos ontológicos, epistemológicos, o tipo de explicação e o método que será usado para mensurar suas variáveis. Lembre-se que sua nova teoria deve estar pautada em variáveis independentes diferentes daquelas usadas por Tocqueville. Não esqueça de explicar suas escolhas, apresentando os mecanismos causais, em no máximo 15 linhas.

B) Teoria: O aumento da distribuição de renda (X) provoca o surgimento da democracia (Y).

Desenho de pesquisa: Verificação da correlação entre as medidas P90/P10 (X) para mensuração da distribuição de renda, e da Polity IV (Y) para mensuração da democracia. Estudo observacional de corte transversal, em que será levantado e analisados dados entre países diferentes para um determinado ano. Para isso, coleta dos dados disponíveis pelas entidades que fazem essas mensurações, como o *Center for Systemic Peace* (CSP) e o *Human Development Report Office for the United Nations Development Programme* (UNDP). A teoria será verificada se houver forte evidência de correlação entre as variáveis, em que um maior valor de X também corresponda com um maior valor de Y.

Explicação : A distribuição de renda, ou seja, a repartição de bens e riqueza em uma nação define a condição de seus indivíduos como agentes da organização pública. É a partir de uma distribuição de renda mais igualitária que os indivíduos de uma nação ampliam suas condições de participação no poder público. Isso se explica pela necessidade, enquanto indivíduo, de riqueza e de bens para a compreensão da estrutura de poder em que ele está inserido. E também é a partir dele que o indivíduo pode interferir-lo. Em outras palavras, a distribuição de renda determina a condição de estabelecimento da democracia.

**Tópicos da Aula 2**

Em um estudo sobre democratização em 28 ex-países socialistas, Steven Fish (1998)1 identificou uma alta correlação entre nível de reformas econômicas e nível de democracia. Para o autor, a realização de reformas econômicas logo após o fim do regime socialista (privatização, liberalização de preços, menos intervenção do Estado na economia) possibilitou um maior nível de democratização. Logo, países que realizaram poucas reformas econômicas em direção a uma economia capitalista, democratizaram-se menos. A relação causal é explicada por meio do seguinte mecanismo: reformas econômicas possibilitam o surgimento de classes médias e elites econômicas que, por dependerem menos de empregos estatais como fonte renda, são mais livres para demandarem maior participação e abertura política.

Com base apenas no enunciado, no gráfico e nas quatro etapas de análise da causalidade, responda:

(A) Há mecanismo causal crível que conecta X a Y? Se não há, justifique (máximo de 3 linhas).

A) Sim. O eventual sucesso de reformas econômicas podem emergir novas classes econômicas que disponham de mais interesses e capacidade para atuação política e, assim, há um aumento na demanda para abertura política daquele país.

(B) Podemos eliminar a possibilidade de que Y causa X? Se não podemos, justifique (máximo de 3 linhas).

B) Sim. Como o estudo foi feito com países que saíram de regimes socialistas, não seria possível que um país fosse democrático e em decorrência disso as reformas teriam sido realizadas. Assim o fator temporal afasta essa possibilidade.

(C) Existe covariação entre X e Y? Se não há, justifique (máximo de 3 linhas).

Sim. Pelo gráfico é possível notar a correlação entre as variáveis. Quanto maior o score da reforma econômica, há mais observações de países com o rating da reforma democrática baixo e vice-versa.

(D) Controlou-se por todas as variáveis colineares Z que podem tornar a associação entre X e Y espúria? Se não controlou-se, justifique (máximo de 5 linhas).

Não. Pelo excerto do estudo não é possível constar que o pesquisador isolou variáveis que poderiam influenciar no fenômeno estudado. Por exemplo, é possível que os países que mais se democratizaram possuíssem uma religião majoritária mais propícia para essa forma de governo.

(E) Trata-se de um experimento ou de um estudo observacional? Justifique (máximo de 3 linhas).

Trata-se de um estudo observacional, pois o pesquisador utilizou-se de dados disponíveis no mundo para o atestamento de sua teoria. Além disso, ele não tinha a capacidade de controle sobre variável independe para realizar um experimento para esse estudo.

**Tópicos da Aula 3**

Para responder os exercícios aplicados da Aula 3, baixe a base de dados (*CPDS\_2016.xlsx*) e seu *Codebook* (*Codebook.pdf*), disponíveis no Moodle. Ela é composta por algumas variáveis da “*Comparative Political Data Set*” para o ano de 2016.

Antes de iniciar a resolução, dê uma olhada no banco de dados e no *codebook* para identificar o os conceitos que cada uma das variáveis representa e como foram operacionalizadas para a análise empírica. Esse processo é muito importante para saber se as medidas são confiáveis e válidas, além de saber como proceder nossa análise desde as estatísticas e dados descritivos. Feito isso, responda às seguintes questões:

(A) As variáveis “*bic*” e “*judreview*” são variáveis que operacionalizam características institucionais. Existe alguma diferença na operacionalização das variáveis? Justifique.

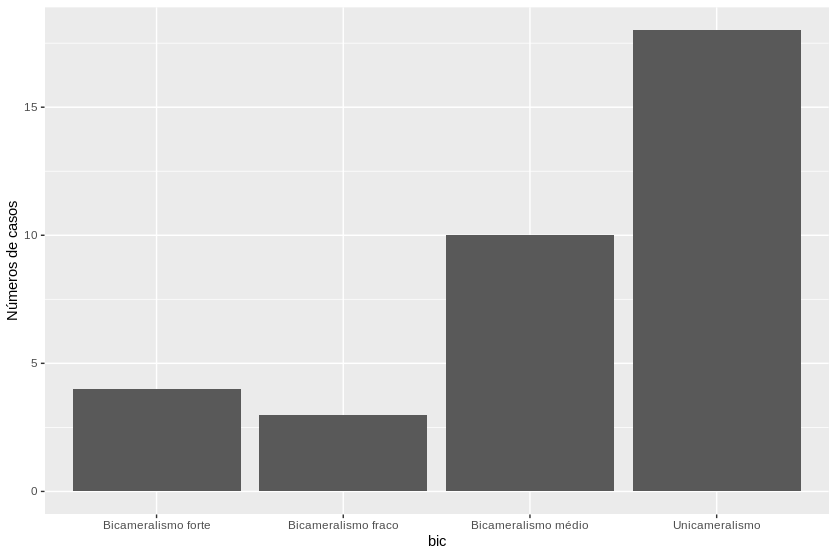
A) Sim. A variável “bic” pode ser descrita como do tipo ordinal, pois ela descreve o nível de bicameralismo, ou seja, ela pode ser ordenada. Já a variável “judreview” apenas pode assumir dois valores, sim ou não, que não podem ser ordenados entre sim. Assim, ela é pode ser classificada como do tipo categórica.

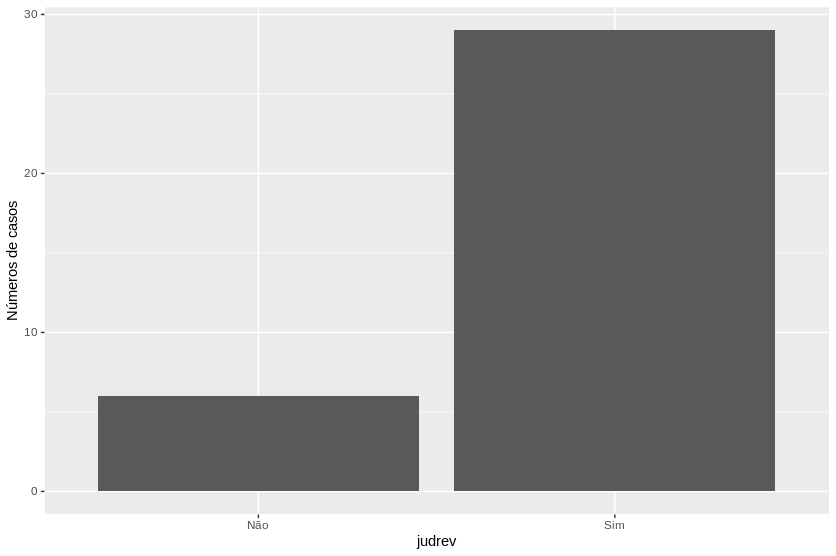
(B) Qual a diferença da variável “*womenpar*”, da participação de mulheres no parlamento, com as duas variáveis do exercício A?

B) A variável “womanpar”, ao contrário das outras duas variáveis, tem como característica haver diferença de unidade constante nos valores que pode assumir, pois representa uma porcentagem. Assim, é classificada como do tipo contínua.

(C) Faça a representação gráfica mais adequada para representar as variáveis “*bic*” e “*judreview*”. Justifique suas escolhas.

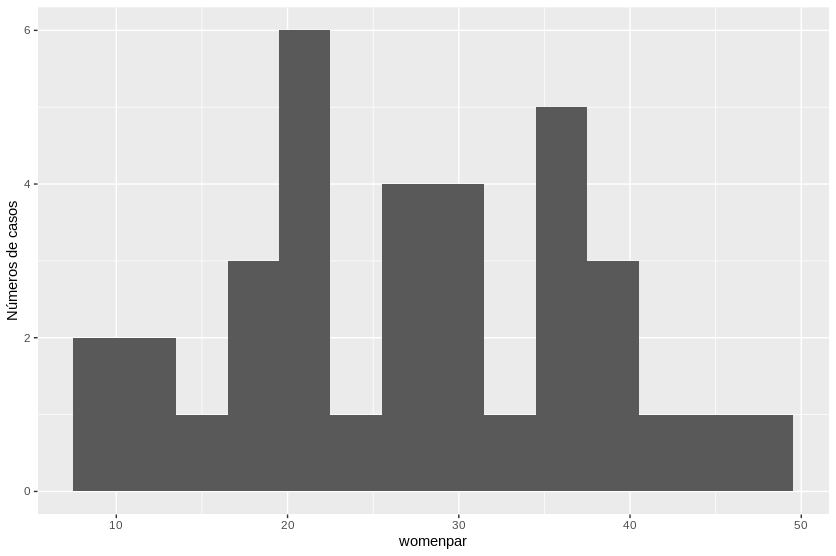
C) O tipo de gráfico mais adequado para a representação de variáveis categóricas é o gráfico de barras, visto que o gráfico de setores é muito criticado por especialistas.





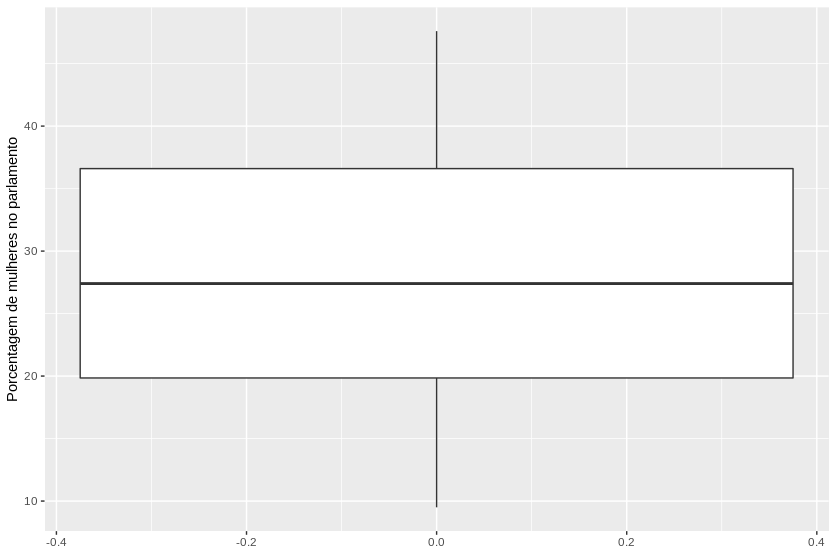
(D) Faça a representação gráfica mais indicada para apresentar a distribuição da variável “*womenpar*”. Justifique a escolha.

D) O histograma é um gráfico que mostra a distribuição de uma variável contínua, tendo como o eixo x os valores da variável, e o eixo y a frequência relativa das observações. Definindo uma quantidade razoável de blocos a ser exibida, o gráfico transmite bem a distribuição.



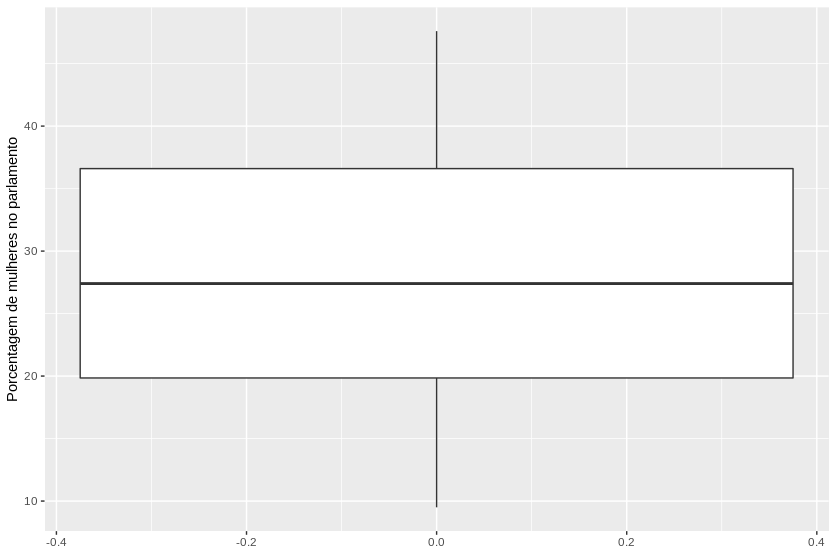
(E) A variável “*womenpar*” possui *outliers* (valores atípicos)? Justifique sua resposta graficamente.

E) Não. O gráfico box-plot não indica valores discrepantes.



(F) A distribuição das observações na variável “*womenpar*” é simétrica ou assimétrica? Justifique graficamente sua resposta.

F) É assimétrica. O gráfico box-plot mostra a linha mediana mais próxima ao primeiro percentil, o que indica que os dados são assimétricos positivos



(G) Qual medida de tendência central é mais adequada para cada uma das três variáveis? Por que? Apresente as medidas e como foram obtidas.

G) A única medida da tendência central apropriada para variáveis categóricas é a moda, assim, ela é a única possível para as variáveis “bic” e “judreview”. Para a variavél contínua “womenpar” a medida mediana é mais adequada, pois os valores extremos não a influenciam. As modas das variáveis categóricas foram obtidas pelas observações dos gráficos em barra. O valor da mediana da variável womenpar foi obtida usando a função median() da linguagem de programação R.

bic: Moda = 1 (unicameralimso)

judreview: Moda = 1 (Sim)

womenpar: Mediana = 27.4

(H) Calcule a variância e o desvio-padrão e explique o que eles indicam.

H)

Média de womenpar (x̄) =

(28,7+31,1+38+20,4+26+12,6+17,9+20+37,4+23,8+41,5+26,2+36,5+19,7+10,1+47,6+22,2+31+9,5+18+21,3+28,3+12,9+37,3+31,4+39,6+27,4+34,8+13,7+20+36,7+39,1+43,6+32+19,2) / 35 = 27.3

Variância de womenpar (s²) =

[(28,7-27.3)²+(31,1-27.3)²+(38-27.3)²+(20,4-27.3)²+(26-27.3)²+(12,6-27.3)²+(17,9-27.3)²+(20-27.3)²+(37,4-27.3)²+(23,8-27.3)²+(41,5-27.3)²+(26,2-27.3)²+(36,5-27.3)²+(19,7-27.3)²+(10,1-27.3)²+(47,6-27.3)²+(22,2-27.3)²+(31-27.3)²+(9,5-27.3)²+(18-27.3)²+(21,3-27.3)²+(28,3-27.3)²+(12,9-27.3)²+(37,3-27.3)²+(31,4-27.3)²+(39,6-27.3)²+(27,4-27.3)²+(34,8-27.3)²+(13,7-27.3)²+(20-27.3)²+(36,7-27.3)²+(39,1-27.3)²+(43,6-27.3)²+(32-27.3)²+(19,2-27.3)²] / 35 - 1

= 104.0106

Desvio-padrão de womanpar () = = 10,19856

A variância e o desvio-padrão indicam o quanto a porcentagem de mulheres no parlamento, das observações, variam acerca da média. Assim, pelos dados que obtivemos, podemos constar que, dada as observações da planilha, a porcentagem de mulheres no parlamento varia em um grau médio.

Script para realização de exercícios na linguagem R

|  |
| --- |
| **library**(tidyverse) **library**(readxl)  *#Tópicos da Aula 3* *#Exercício C* CPDS\_2016 <- read\_excel("Lista 4/CPDS\_2016.xlsx") *# Gera gráfico de barras para a variável bic* CPDS\_2016 %>%   ggplot(aes(x = bic)) +   geom\_bar() +  labs(y = "Números de casos") *#Gera gráfico de barras para a variável judreview* CPDS\_2016 %>%   ggplot(aes(x = judrev)) +   geom\_bar() +  labs(y = "Números de casos")  *#Exercício D* *#Gera histograma para a variável womenpar* CPDS\_2016 %>%   ggplot(aes(x = womenpar)) +   geom\_histogram(binwidth = 3) +  labs(y = "Números de casos")  *#Exercício E e F* *#Gera gráfico o box-plot para a variável womenpar* CPDS\_2016 %>%   ggplot(aes(y = womenpar)) +   geom\_boxplot() +  labs(y = "Porcentagem de mulheres no parlamento")  *#Exercício H* *#Obtendo a variância de womenpar* var(CPDS\_2016$womenpar)  *#Obtendo o desvio-padrão de womenpar* sd(CPDS\_2016$womenpar) |