

Trabalho de Computação em Estatística



Universidade de Brasília (UnB)

Relatório sobre os bilionários

Estudantes: Matheus Donito, João Pedro, Lara Silva

Professor: Guilherme Souza Rodrigues

Brasília, DF

14 de Dezembro, 2023

Sumário

Introdução	3
Objetivo	4
Objetivo Geral	4
Objetivo específico	4
Análise Descritiva	5
Variáveis quantitativas	5
Variáveis qualitativas	5
Desenvolvimento	6
Características populacionais	6
Sexo	6
Distribuição do sexo na amostra	6
Figura 1 - Distribuição do gênero	6
Tabela 1 - Distribuição do sexo entre as categorias	7
Distribuição do sexo por categoria	7
Falta uma tabela	7
Figura 2 - Frequência de bilionários por categoria e gênero	8
Origem da fortuna (Self-made ou herdeiro)	8
Distribuição da variável na amostra	8
Figura 3 - Frequência absoluta Self-Made	9
Figura 4 - Frequência relativa Self-Made	10
Origem da fortuna por gênero	10
Figura 5 - Frequência relativa do Gênero por Self-Made	11
Tabela 2 - Distribuição dos herdeiros por gênero.	11
Tabela 3 - Distribuição dos herdeiros por gênero.	11
Origem da fortuna por Indústria	11
Figura 6 - Proporção de Self-Made por categoria	12
Idade	12
Idade por gênero	12
Figura 7 - Distribuição da idade dos bilionários por gênero	13
Distribuição da idade em Self-Made	13
Figura 8 - Distribuição da idade dos bilionários pela origem da fortuna	14
Correlação entre idade e patrimônio	14
Figura 9 - Relação entre idade e patrimônio	15

Tabela 4 - Coeficiente de Spearman	15
Categoria / Indústria	15
Distribuição dos bilionários e do patrimônio por indústria	15
Figura 10 - Distribuição dos bilionários por Indústria	16
Correlação entre indicadores de um país e bilionários	16
Relação entre bilionários, PIB e população	16
Figura 11 - Regressão Linear entre PIB, quantidade de bilionários e população	16
Figura 12 - Gráfico de dispersão entre a quantidade de bilionários e o PIB	17
Figura 13 - Gráfico de dispersão entre a população e o PIB	18
Relação entre expectativa de vida e bilionários	18
Figura 14 - Gráfico de dispersão entre a quantidade de bilionários e a expectativa de vida	19
Tabela 5 - Coeficiente de Spearman	19
Relação entre matrículas no ensino superior e bilionários	19
Figura 15 - Gráfico de dispersão entre a quantidade de bilionários e matrículas no Ensino Superior	20
Tabela 6 - Coeficiente de Spearman	20
Países	20
Número de bilionários e proporção dos gêneros	20
Patrimônio entre os países	21
Conclusão	22
Referências	23
Participação dos integrantes	24

Introdução

Esta pesquisa, realizada por Nidula Elgiriye withana, um estudante de ciência de dados, analisou um conjunto abrangente de dados contendo informações de 2640 bilionários. Destaca-se que 71,34% desses indivíduos alcançaram sua condição financeira por mérito próprio (self-made), enquanto 28,66% herdaram suas fortunas. O objetivo central desta investigação foi modelar e compreender os padrões subjacentes à distribuição de riqueza entre esses bilionários, explorando elementos como origem da fortuna, correlações geográficas, faixa etária e diferenças de gênero. Essa análise detalhada busca não apenas mapear a concentração de riqueza, mas também entender as dinâmicas complexas que impulsionam o sucesso financeiro em diferentes contextos.

Objetivo

Objetivo Geral

O presente trabalho busca encontrar a distribuição de riqueza entre diferentes categorias e indústrias, revelando insights cruciais sobre onde está o setor com a maior parte dos bilionários. Além disso, busca-se encontrar, com esse estudo, uma análise econômica e geográfica da riqueza dos bilionários (Produto Interno Bruto - PIB, expectativa de vida e a relação da matrícula no ensino superior com a quantidade de bilionários), se o bilionário é self-made (sua fortuna foi construída, não herdada), a relação entre a idade, gênero e o acúmulo de patrimônio, por fim a diferença de gênero entre bilionários. Ao examinar essa distribuição, é possível compreender melhor as dinâmicas econômicas e as áreas que mais contribuem para a acumulação de riqueza, o que pode ser vital para estratégias de investimento, desenvolvimento econômico e políticas de redistribuição de recursos. Por fim, identifica-se insights sobre como os indicadores econômicos e de ensino contribuem para a riqueza do país quanto para a quantidade de bilionários, o quanto a idade impacta no desenvolvimento do patrimônio.

Objetivo específico

1. Apresentar a quantidade de bilionários e suas características (self-made, setor, gênero e idade) ;
2. Verificar se há correlação entre o bilionário ser self-made sua idade e gênero;
3. Verificar a distribuição de bilionários entre os países;
4. Encontrar a relação entre PIB com a quantidade de bilionários e população;
5. Analisar como a idade impacta no tamanho da fortuna dos bilionários.

Análise Descritiva

Com um banco de dados do Kaggle, na qual contém dados de milhares de bilionários e diversas características sobre cada um, desde variáveis qualitativas, como quantitativas. Utilizando-se da plataforma R e pela interface do RStudio foram construídos gráficos e tabelas para auxiliar a análise.

Variáveis quantitativas

- finalWorth : O patrimônio líquido final do bilionário em dólares americanos.
- age: A idade do bilionário.
- gdp_country : Produto Interno Bruto (PIB) do país do bilionário.
- item gross_tertiary_education_enrollment : Matrícula no ensino superior no país do bilionário.
- item life_expectancy_country : Expectativa de vida no país do bilionário.
- item population_country: População do país do bilionário.

Variáveis qualitativas

- category: a categoria ou setor em que opera o negócio do bilionário.
- selfMade : Indica se o bilionário é self-made (Verdadeiro/Falso).
- item gender: O gênero do bilionário.

Desenvolvimento

Esta pesquisa, realizada por Nidula Elgiriye withana, um estudante de ciência de dados, analisou um conjunto abrangente de dados contendo informações de 2640 bilionários. Destaca-se que 71,34% desses indivíduos alcançaram sua condição financeira por mérito próprio, enquanto 28,66% herdaram suas fortunas. O objetivo central desta investigação foi modelar e compreender os padrões subjacentes à distribuição de riqueza entre esses bilionários, explorando elementos como origem da fortuna, correlações geográficas, faixa etária e diferenças de gênero. Essa análise detalhada busca não apenas mapear a concentração de riqueza, mas também entender as dinâmicas complexas que impulsionam o sucesso financeiro em diferentes contextos.

Características populacionais

Sexo

Distribuição do sexo na amostra

A amostra é composta por 2640 bilionários, sendo 337 mulheres e 2303 homens. É interessante apontar que pouco mais de 87% da amostra é masculina. Isso evidencia que existe uma desigualdade de gênero quando observa-se a alta classe mundial.

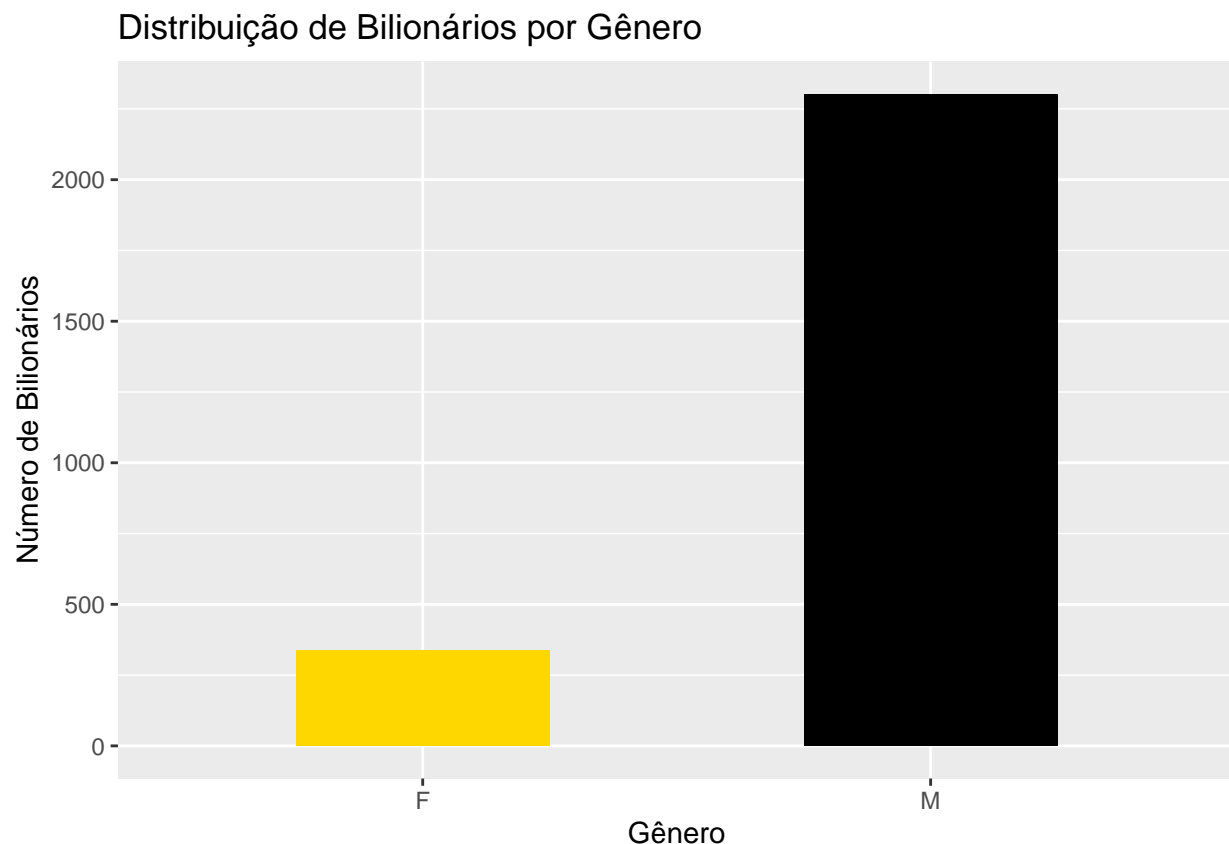


Figura 1 - Distribuição do gênero

	Mulher	Homem	Proporção Homem	Proporção Mulher
Automotive	66	7	0.9041096	0.0958904
Construction & Engineering	39	6	0.8666667	0.1333333
Diversified	158	29	0.8449198	0.1550802
Energy	90	10	0.9000000	0.1000000
Fashion & Retail	227	39	0.8533835	0.1466165
Finance & Investments	342	30	0.9193548	0.0806452
Food & Beverage	167	45	0.7877358	0.2122642
Gambling & Casinos	21	4	0.8400000	0.1600000
Healthcare	173	28	0.8606965	0.1393035
Logistics	34	6	0.8500000	0.1500000
Manufacturing	279	45	0.8611111	0.1388889
Media & Entertainment	79	12	0.8681319	0.1318681
Metals & Mining	65	9	0.8783784	0.1216216
Real Estate	171	22	0.8860104	0.1139896
Service	44	9	0.8301887	0.1698113
Sports	33	6	0.8461538	0.1538462
Technology	284	30	0.9044586	0.0955414
Telecom	31	0	1.0000000	0.0000000

Tabela 1 - Distribuição do sexo entre as categorias

Distribuição do sexo por categoria

Tendo como base a tabela (Tabela 1) e o gráfico (Figura 2) acima e a distribuição de mulheres na amostra, é possível notar que existe uma concentração masculina nos setores, sendo o setor com maior proporção de mulheres o de comida e bebida (Food & Beverage), com 21,23%. Isso mostra como existe uma hegemonia masculina na indústria.

Falta uma tabela

```
## 'summarise()' has grouped output by 'category'. You can override using the
## '.groups' argument.
```

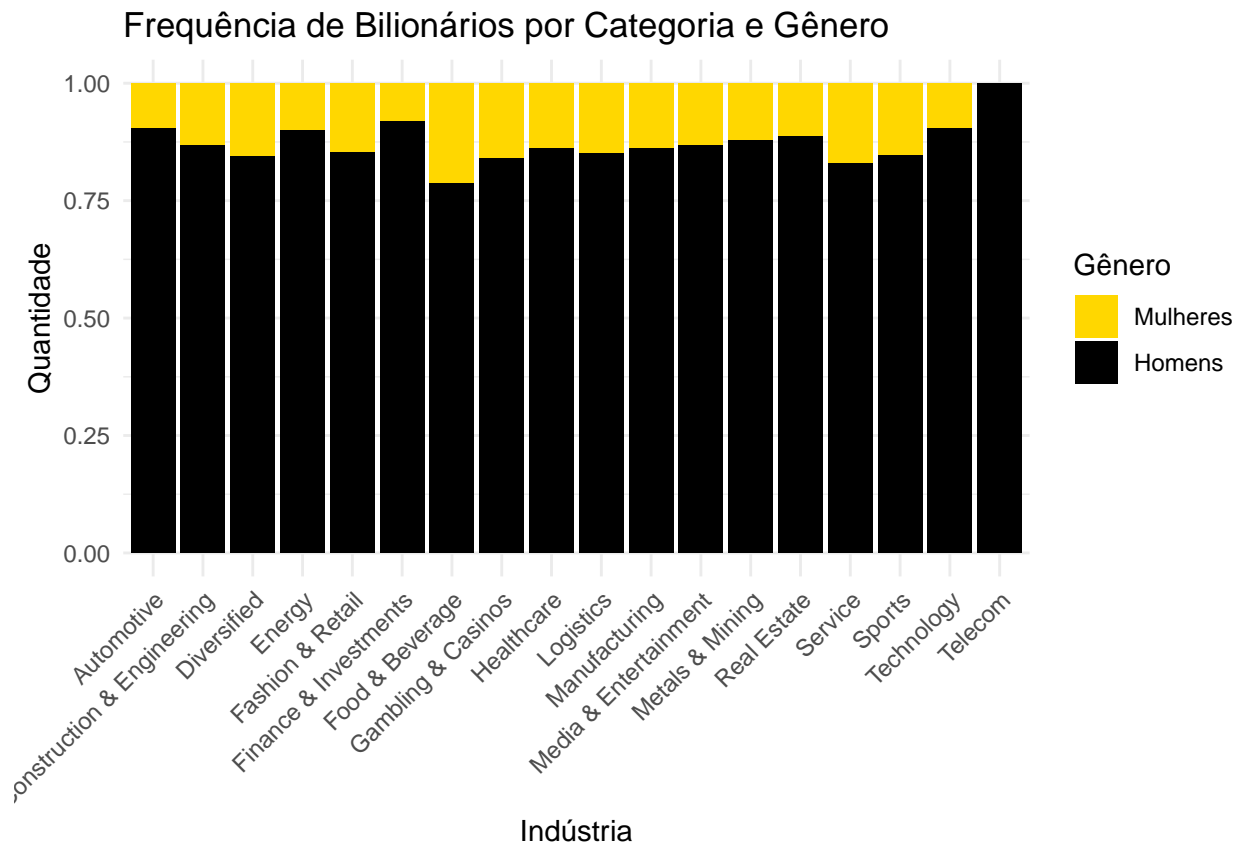



Figura 2 - Frequência de bilionários por categoria e gênero

Origem da fortuna (Self-made ou herdeiro)

Distribuição da variável na amostra

Há 828 bilionários herdeiros e 1812 bilionários self-made. Dessa forma, a maioria dos bilionários presentes no banco de dados são self-made (68,64%), mais que o dobro da quantidade de bilionários herdeiros.

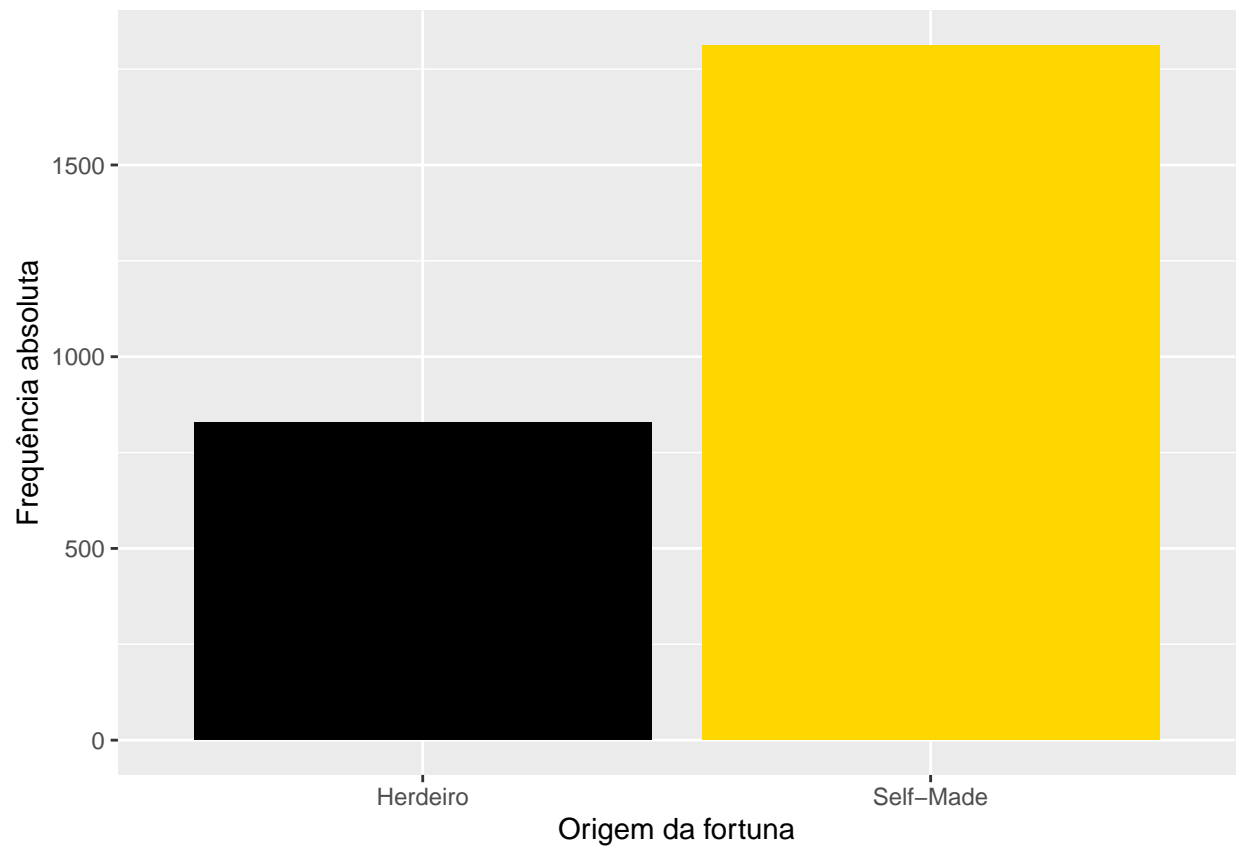


Figura 3 - Frequência absoluta Self-Made

```
## Don't know how to automatically pick scale for object of type <table>.  
## Defaulting to continuous.
```

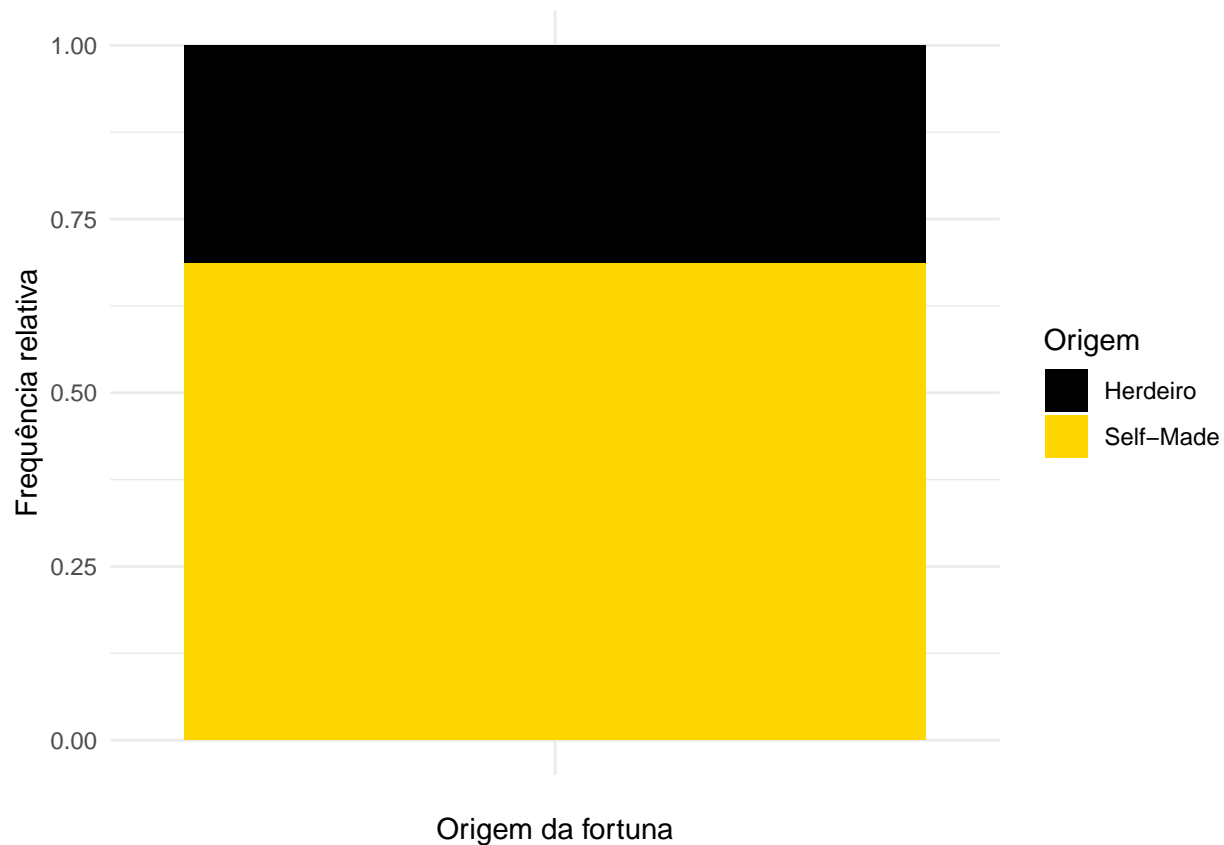


Figura 4 - Frequência relativa Self-Made

Origem da fortuna por gênero

Pela tabela, é possível observar que a proporção de bilionárias herdeiras mulheres é de 71%, logo, é discrepante em relação à proporção de bilionários herdeiros homens, apenas 25%. Dessa forma, dentre todos os 2640 bilionários presentes na amostra, apenas 96 são mulheres selfmade. Além disso, é possível dizer que existe uma associação mediana entre o gênero e o fato da pessoa ser herdeira ou não, uma vez que o Coeficiente de Contingência foi de 0,44.

	Herdeiro	Self-Made
Homem	25.48849	74.51151
Mulher	71.51335	28.48665

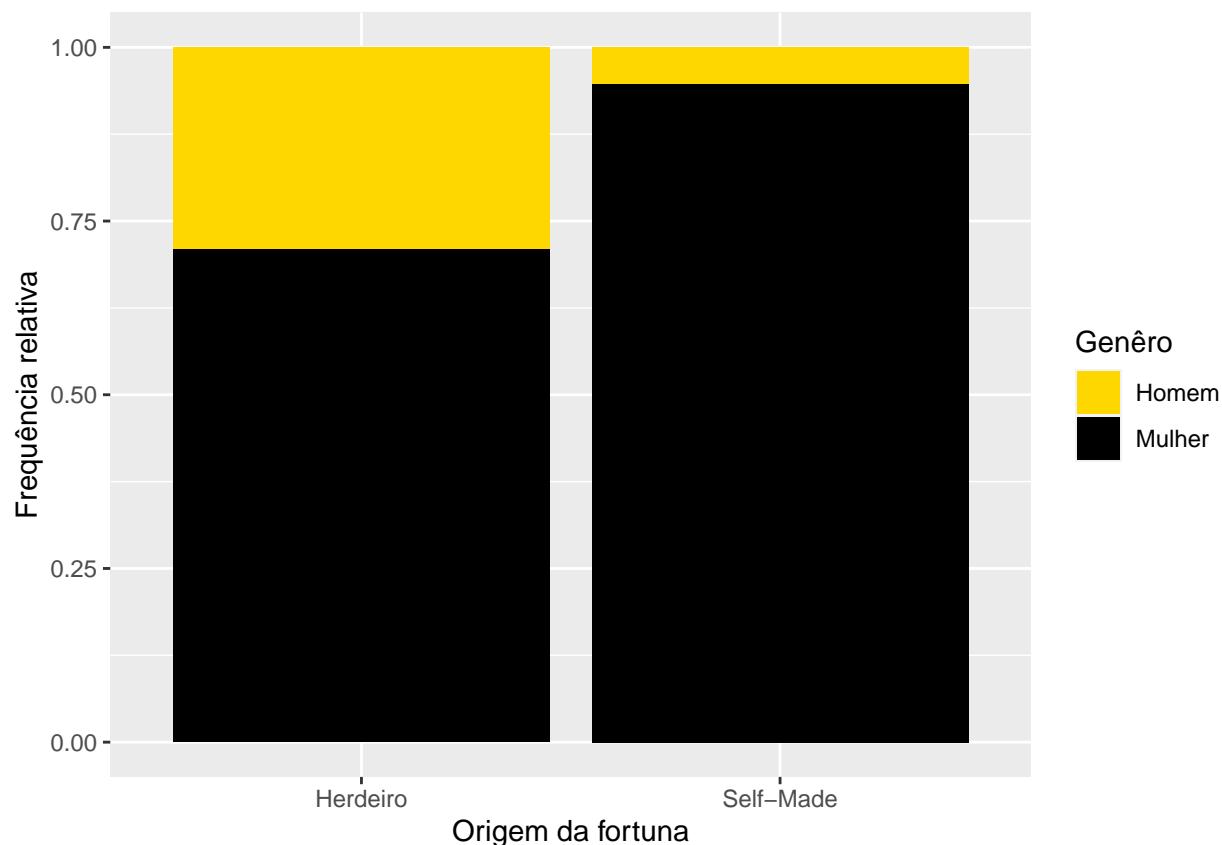


Figura 5 - Frequência relativa do Gênero por Self-Made

	Herdeiro	Self-Made
Homem	25.48849	74.51151
Mulher	71.51335	28.48665

Tabela 2 - Distribuição dos herdeiros por gênero.

Coeficiente.de.Spearman
0,44

Tabela 3 - Distribuição dos herdeiros por gênero.

Origem da fortuna por Indústria

Analisando o gráfico e a tabela acima, é possível notar que apenas uma categoria de indústria possui maioria em bilionários herdeiros (diversas), com 55,61% como frequência relativa. E que todas as outras possuem bilionários self-made como maioria, com destaque para a indústria de tecnologia, com 92,99% e telecomunicações, com 83,87%.

A distribuição de bilionários self-made por setor revela uma disparidade significativa em diversas indústrias. Enquanto setores como tecnologia e telecomunicações demonstram uma presença dominante de indivíduos que construíram suas fortunas de forma independente, com 92,99% e 83,87% dos bilionários sendo self-made, respectivamente. Enquanto outros campos, como construção e alimentação, exibem uma presença comparativamente menor de bilionários que construíram suas próprias fortunas, com 48,89% e 48,58% dos bilionários sendo herdeiros, respectivamente, têm uma distribuição mais equilibrada entre aqueles que construíram suas fortunas por conta própria e aqueles que as herdaram. Essa variedade na distribuição sugere uma intersecção complexa entre os setores e os caminhos para o status bilionário.

```
## 'summarise()' has grouped output by 'category'. You can override using the
## '.groups' argument.
```

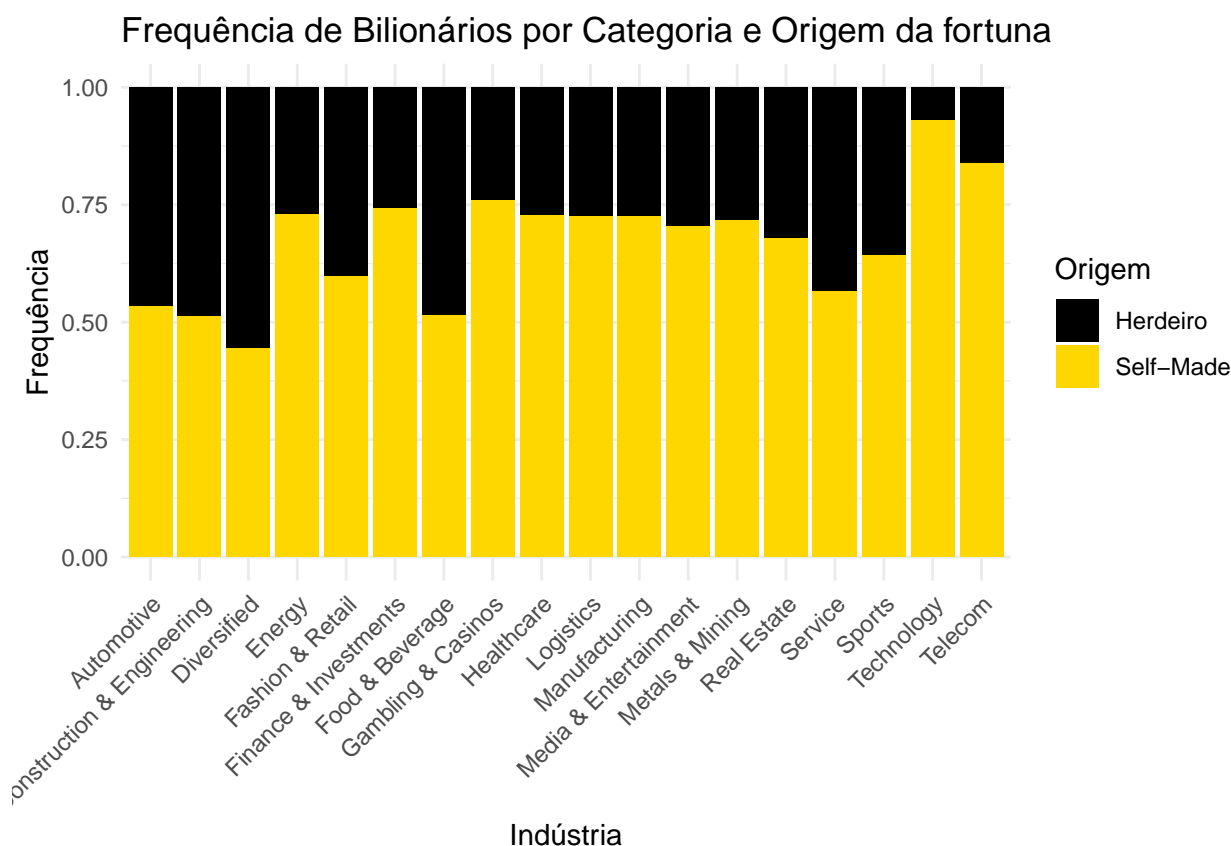


Figura 6 - Proporção de Self-Made por categoria

Idade

Idade por gênero

Analisando o gráfico “Distribuição da idade dos bilionários por gênero” e a tabela “Características da idade por sexo”, é perceptível que os homens bilionários são ligeiramente mais velhos que as mulheres bilionárias. Apesar de que há mais outliers de homens com idade abaixo do limite inferior do boxplot do que de mulheres. É interessante apontar, além de tudo isso, que as idades entre os gêneros são praticamente as mesmas.. Isso fica claro quando a tabela evidencia que a

idade mediana dos homens 65 e das mulheres 63, e ambos os gêneros com desvio padrão parecidos.

Distribuição da Idade dos Bilionários por Gênero

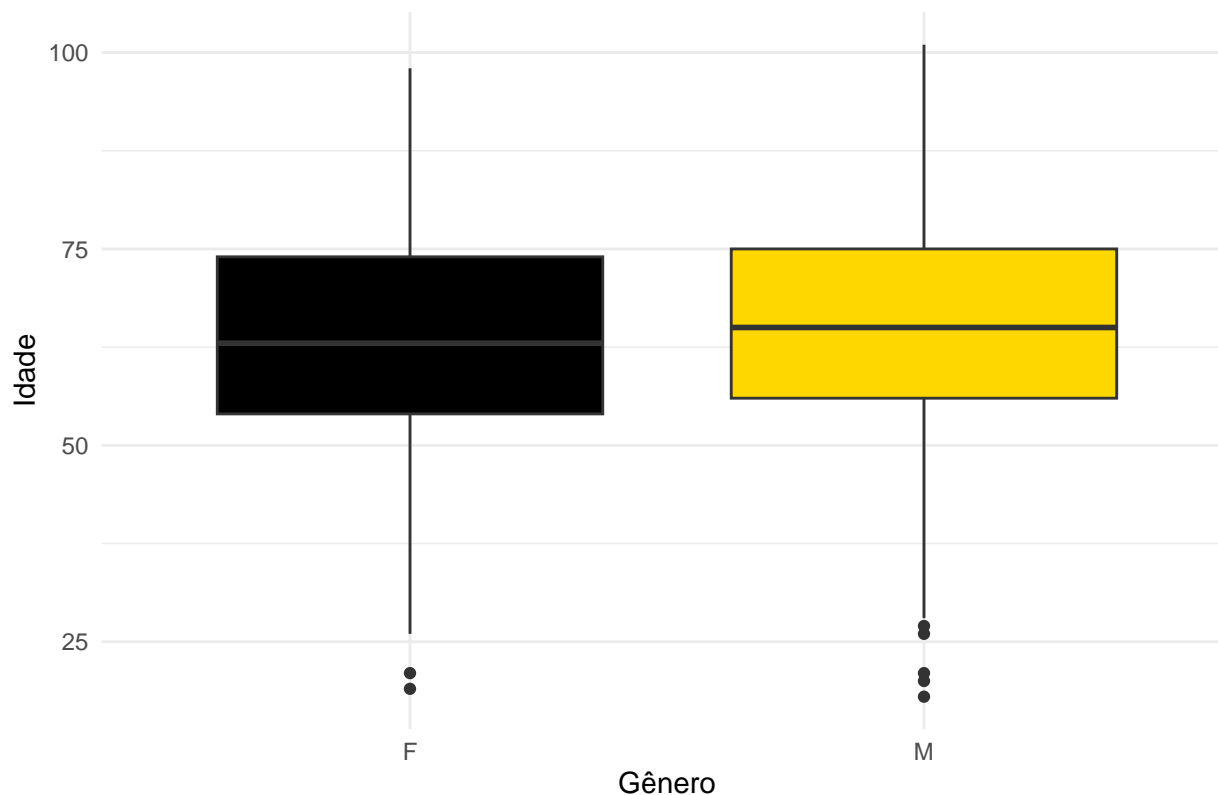


Figura 7 - Distribuição da idade dos bilionários por gênero

Distribuição da idade em Self-Made

Analisando o gráfico “Distribuição da idade dos bilionários pela origem da fortuna”, é perceptível que os bilionários self-made são ligeiramente mais novos que os bilionários herdeiros. Pela tabela é visível que a idade não varia muito entre herdeiros e não herdeiros, possuindo uma mediana de 68 e 64 respectivamente, com desvio padrão parecido entre os dois grupos. Apesar disso, os bilionários herdeiros compõem a maior parte dos bilionários menores que 25 tendo uma presença maior de outliers do que em relação aos bilionários self-made. A medida de associação R quadrado entre idade e self-made é 0,002 muito fraca ou quase nula.

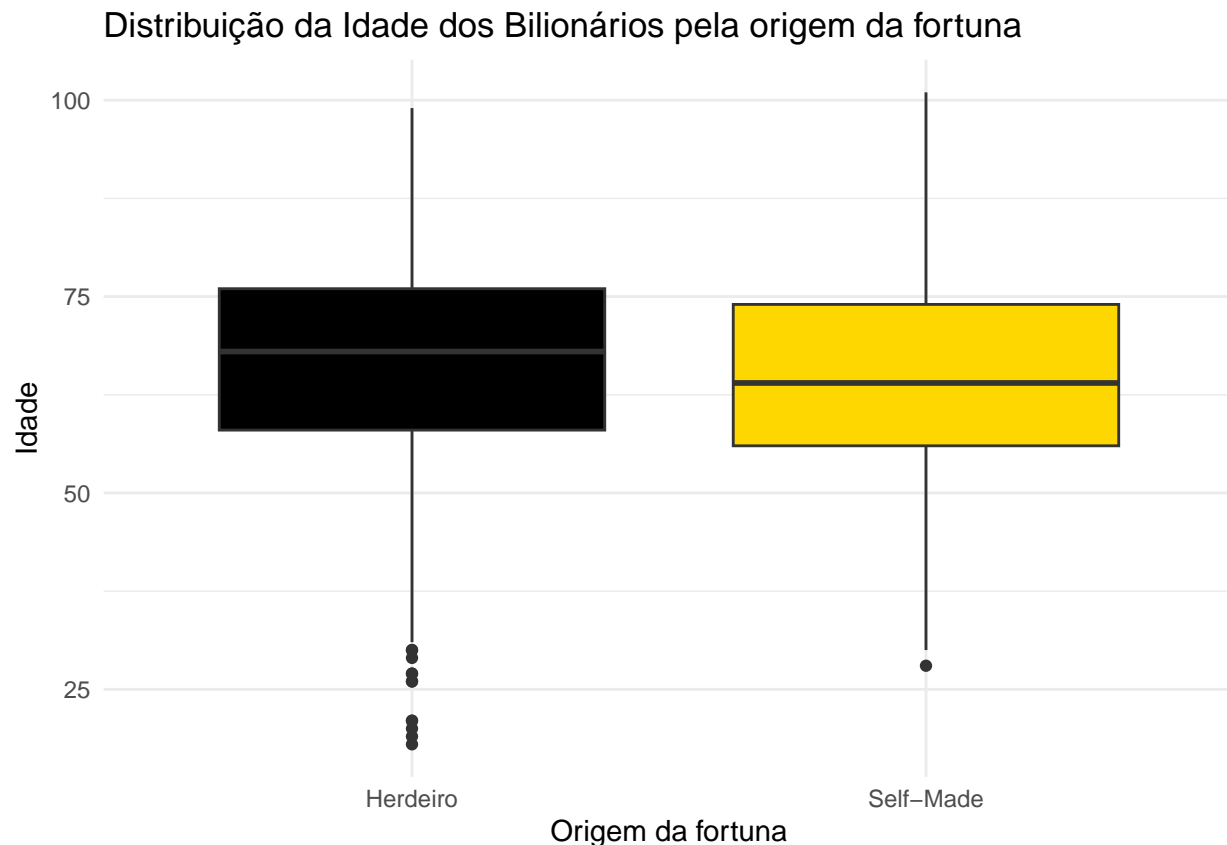


Figura 8 - Distribuição da idade dos bilionários pela origem da fortuna

Correlação entre idade e patrimônio

Ao analisar a relação entre a idade e o patrimônio líquido dos bilionários, a correlação de Spearman de 14,40% sugere uma associação positiva, mas não linear, entre essas variáveis. Em contraste com a correlação de Pearson, que mede a força e direção de uma relação linear, a correlação de Spearman avalia a relação monotônica, ou seja, se as variáveis tendem a aumentar ou diminuir juntas, porém não necessariamente a uma taxa constante.

A correlação de 14,40% indica que, à medida que a idade dos bilionários aumenta, há uma tendência geral de que o patrimônio líquido também aumente, mas essa relação não é linear. Em outras palavras, a idade e o patrimônio líquido dos bilionários têm uma associação positiva, mas as mudanças em uma variável não são preditivas de mudanças proporcionais na outra.

É crucial ressaltar que a correlação de Spearman de 14,40% indica uma associação fraca entre idade e patrimônio líquido. Isso sugere que outros fatores além da idade têm um impacto substancial no patrimônio líquido dos bilionários. É possível que outros aspectos, como educação, oportunidades de negócios, investimentos estratégicos, entre outros, desempenhem papéis mais significativos na determinação do patrimônio líquido, superando a influência da idade. Portanto, embora haja uma associação positiva entre idade e patrimônio líquido, é importante reconhecer que essa relação não é linear e que outros fatores também desempenham um papel crucial na determinação do patrimônio líquido dos bilionários.

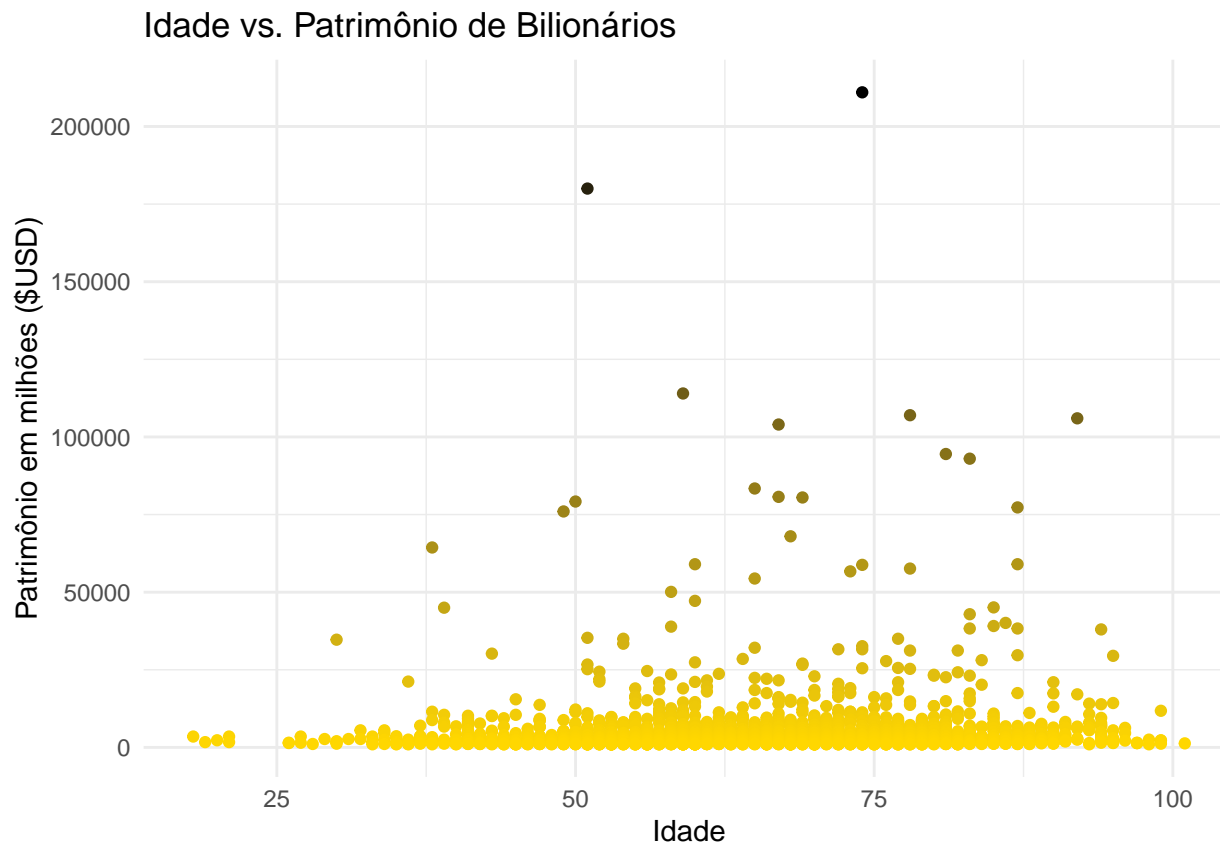


Figura 9 - Relação entre idade e patrimônio

Coefficiente.de.Spearman
0.1440389

Tabela 4 - Coeficiente de Spearman

Categoria / Indústria

Distribuição dos bilionários e do patrimônio por indústria

Olhando para a tabela “Distribuição dos bilionários por Indústria” e para o gráfico (Figura 9), é possível notar que a área de finanças e investimentos abriga a maior proporção de bilionários, totalizando 14,09% de todos os bilionários na amostra. Em seguida, estão manufatura, tecnologia e moda, representando, respectivamente, 12,27%, 11,89% e 10,08%.

É interessante observar, além do que foi mencionado anteriormente, que existe uma diferença discrepante na quantidade de bilionários por setor. Enquanto a indústria de finanças e investimentos contém 372 bilionários, os setores de construção e engenharia, jogos de azar e cassinos, logísticas, esportes e telecomunicação apresentam menos de 50. Dessa forma, finanças e investimentos, além de representar 14,09% dos bilionários, ele sozinho contém mais que o dobro da quantidade dos 5 setores citados anteriormente.

No entanto, em termos de patrimônio total, as indústrias líderes são tecnologia, moda, finanças e alimentação, juntas compreendendo 54% do patrimônio total dos bilionários nessas áreas da amostra. Isso é evidenciado no gráfico a seguir.

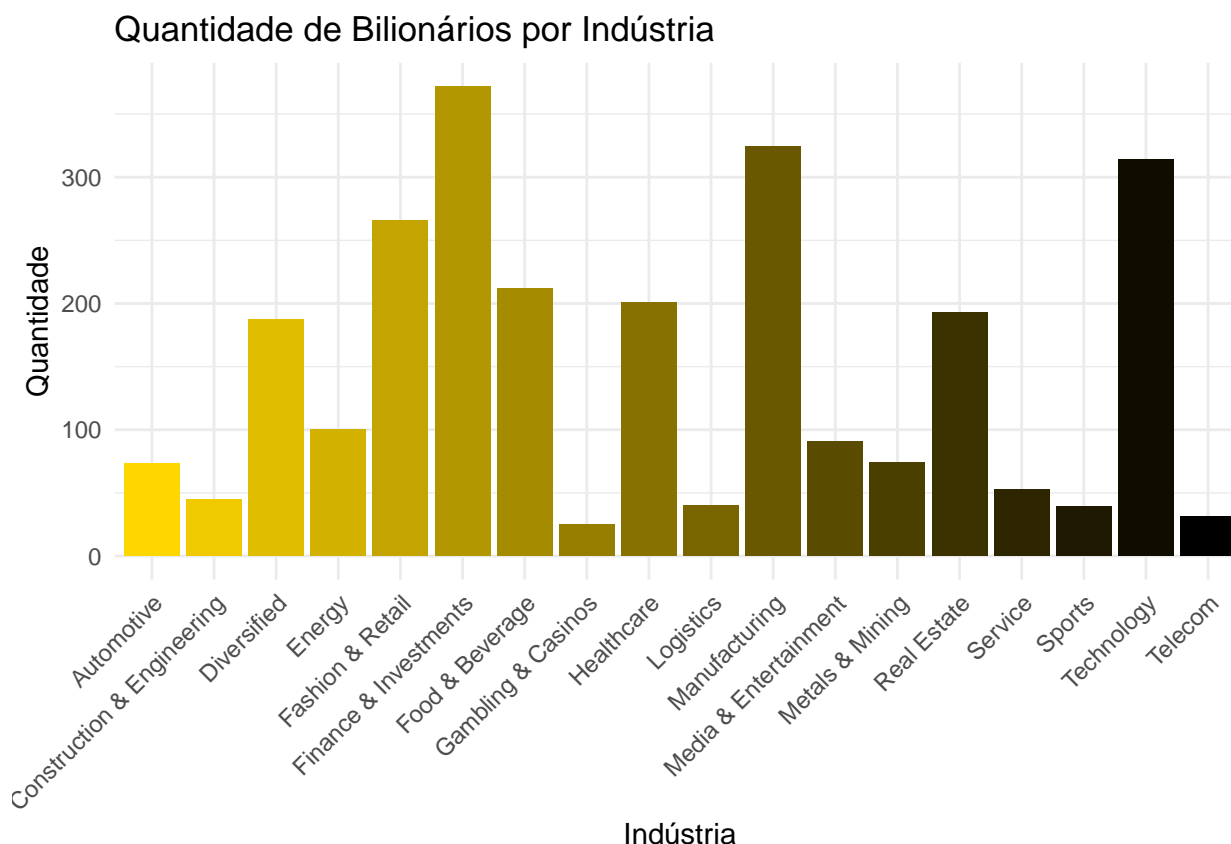


Figura 10 - Distribuição dos bilionários por Indústria

Correlação entre indicadores de um país e bilionários

Relação entre bilionários, PIB e população

A partir dos gráficos de dispersão acima pode-se observar a relação linear entre PIB e a quantidade de bilionários, com ele é possível inferir uma correlação positiva, ou seja, quanto maior a quantidade de bilionários maior será o PIB do país. Além disso, no outro gráfico é possível enxergar uma certa correlação linear, para verificarmos a relação dessas duas variáveis no PIB se torna interessante um modelo de regressão linear.

$$PIB = \beta_0 + \beta_1 \text{ quantidade_bilionários} + \beta_2 \text{ população_país}$$

$$PIB = 113129875908,8 + 29930580898,9 * \text{quantidade_bilionários} + 772.7 * \text{população_país}$$

$$R^2 = 0,94$$

Figura 11 - Regressão Linear entre PIB, quantidade de bilionários e população

O intercepto representa o valor esperado do PIB quando todas as variáveis independentes (quantidade de bilionários e população do país) são zero. Neste caso, sugere-se que, mesmo na ausência de bilionários e

população, o PIB seria aproximadamente U\$113,13 bilhões. Porém, na prática, este valor pode não ter um significado prático real.

Com relação, aos coeficientes das variáveis independentes referem-se a quanto o PIB varia a cada unidade a mais dessa variável. O coeficiente da quantidade de bilionários sugere a mudança esperada no PIB próximo de U\$29,93 bilhões quando a quantidade de bilionários aumenta em uma unidade, mantendo a população do país constante. Enquanto, o coeficiente da população do país indica a mudança esperada no PIB de aproximadamente U\$772,7 quando a população do país aumenta em uma unidade, mantendo a quantidade de bilionários constante.

Considerando um conjunto de dados que contempla 44 bilionários e uma população de 212 milhões, a projeção do Produto Interno Bruto (PIB) para o Brasil é estimada em aproximadamente U\$1,59 trilhões, conforme um modelo que considera a relação entre a quantidade de bilionários, a população e o PIB. No entanto, o valor real do PIB apresentado é de U\$1,84 trilhões. Essa discrepância sugere uma precisão de 86% para modelo e o valor real observado, indicando possíveis variáveis ou fatores não considerados na modelagem que certamente influenciam a economia do país.

```
## 'geom_smooth()' using formula = 'y ~ x'
```

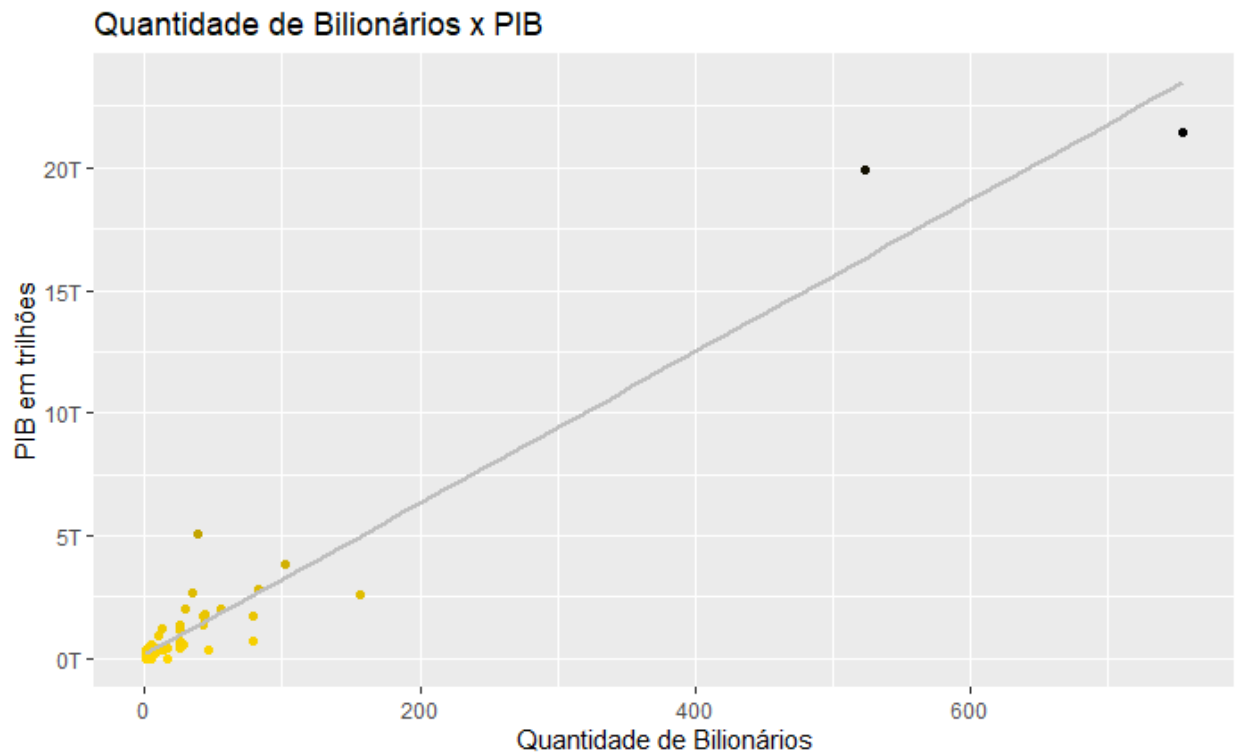


Figura 12 - Gráfico de dispersão entre a quantidade de bilionários e o PIB

```
## 'geom_smooth()' using formula = 'y ~ x'
```

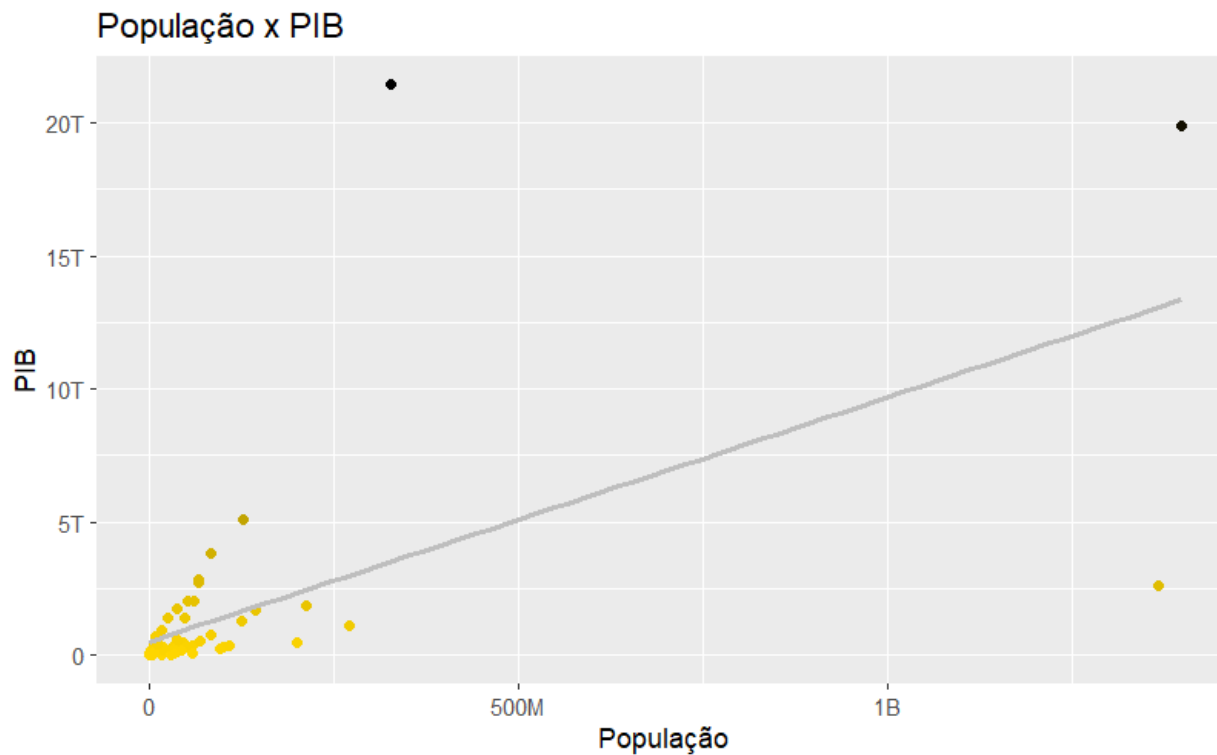


Figura 13 - Gráfico de dispersão entre a população e o PIB

Relação entre expectativa de vida e bilionários

Embora haja uma certa correlação positiva entre o número de bilionários e a expectativa de vida de um país, essa relação é tênue. Como demonstrado no gráfico abaixo, à medida que a expectativa de vida aumenta, também aumenta a quantidade de bilionários.

Os pontos dispersos formam uma nuvem irregular, sugerindo a ausência de uma relação linear clara entre as variáveis. No entanto, através da correlação de Spearman, é possível identificar uma correlação positiva e fraca de 0,319. Isso indica que, embora exista uma correlação, ela não é forte. Gráfico de dispersão entre expectativa de vida e quantidade de bilionários no país

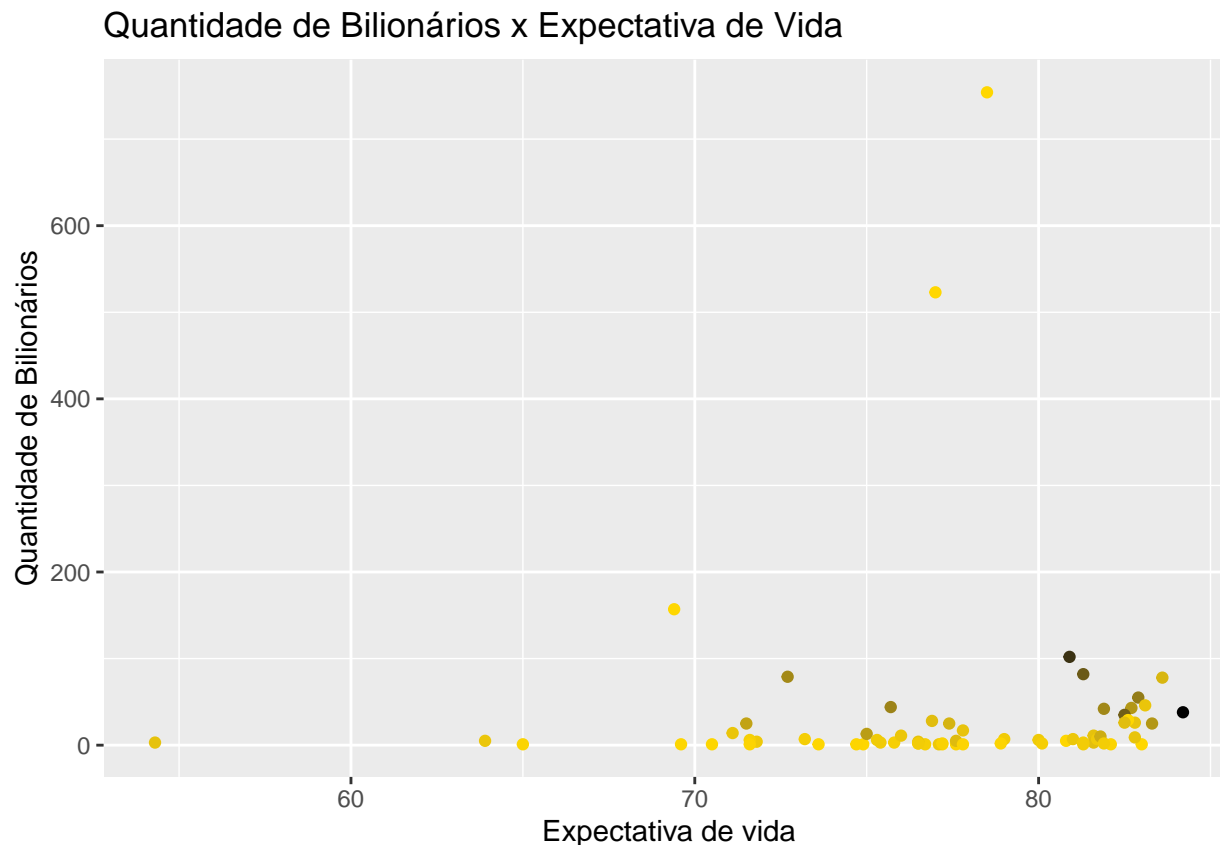


Figura 14 - Gráfico de dispersão entre a quantidade de bilionários e a expectativa de vida

Coefficiente.de.Spearman
0.3187218

Tabela 5 - Coeficiente de Spearman

Relação entre matrículas no ensino superior e bilionários

Assim, como a correlação entre a expectativa de vida e a quantidade de bilionários, o número de matrículas no ensino superior revela insights interessantes sobre como essa variável se interliga com a quantidade de bilionários. É notável que, ao analisar essa dinâmica, seja crucial considerar não apenas a simples associação, mas também a natureza não linear dessas relações.

A utilização do coeficiente de Spearman se faz necessária quando lidamos com relações que não seguem uma forma linear, permitindo uma avaliação mais precisa da associação entre essas variáveis. No entanto, mesmo com essa análise mais complexa, a correlação identificada, de 0,319, aponta para uma conexão positiva e fraca entre esses fatores. Isso significa que, em geral, à medida que a quantidade de matrículas no ensino superior aumenta, a tendência é que o número de bilionários também aumente.

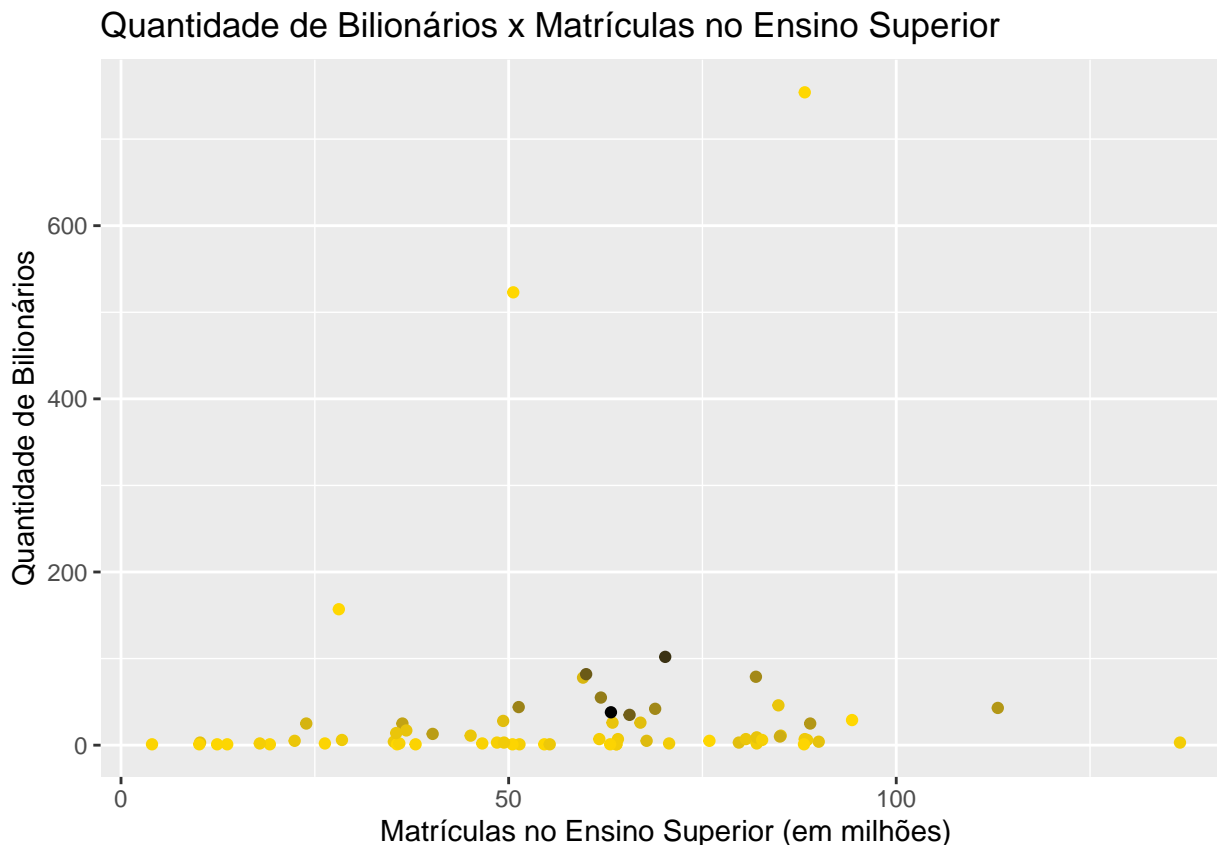


Figura 15 - Gráfico de dispersão entre a quantidade de bilionários e matrículas no Ensino Superior

Coeficiente.de.Spearman
0.3179955

Tabela 6 - Coeficiente de Spearman

Países

Para essa análise, foram considerados os oito países com maior número de bilionários e patrimônio. Para calcular o patrimônio foram somados os valores patrimoniais de todas as indústrias do país.

Número de bilionários e proporção dos gêneros

Tendo como base a tabela acima, é possível notar que existe uma concentração de bilionários nos Estados Unidos e China que, as duas maiores economias do mundo. Se somado o número de bilionários nos dois países, encontra-se um número equivalente a 48% da amostra. Essa diferença se torna ainda mais evidente quando é realizada uma comparação entre o segundo país com maior número de bilionários com o terceiro, onde observa-se que a China possui três vezes mais bilionários que a Índia.

Olhando para os gêneros, é notório que existe uma concentração masculina nos bilionários, sendo o país com maior concentração de mulheres a Alemanha, com 26,47%. O caso mais extremo da amostra é a Rússia, onde mais de 98% dos bilionários são homens.

Patrimônio entre os países

A tabela mostra como a distribuição de patrimônio acontece de maneira desigual entre os países. Somente os Estados Unidos concentram quase 38% de todo o patrimônio da amostra. Quando comparado com a China, a segunda com maior patrimônio, os Estados Unidos apresentam 2,5 vezes mais patrimônio.

Conclusão

Através do banco de dados bilionários disponibilizado pela plataforma Kaggle, realizado por Nidula Elgiriye-withana. Modelamos informações de 2640 bilionários para extrairmos e compreendermos padrões relacionados à distribuição de riqueza entre eles.

Ao destrinchar o banco de dados, percebe-se que há discrepância entre homens (2303) e mulheres (337) na amostra. Essa discrepância é acentuada quando analisamos a influência do gênero em cada setor. Visto que estes possuem pelo menos 78% de homens em sua composição, com destaque para o setor de telecomunicação, que não há representação feminina, seguido pelo setor de investimentos e finanças assim como tecnologia, com apenas 8,06% e 9,55% respectivamente.

O patrimônio dos bilionários tem correlação positiva com a idade dos mesmo, porém é uma relação fraca e não linear dado que o coeficiente de Spearman é 0,144. Com relação a ser self-made e a idade dos bilionários, os herdeiros são ligeiramente mais velhos, há mais bilionários mais novos herdeiros do que self-made, sendo o bilionário herdeiro mais novo de apenas 18 anos, 10 anos mais novo que o bilionário self-made mais novo.

Quanto à idade dos homens e do gênero, apesar de serem ligeiramente mais velhos, há uma quantidade considerável de outliers de homens bilionários com idade inferior ao limite inferior do boxplot em comparação com as bilionárias. Além disso, a quantidade de herdeiras mulheres compõem 71% das bilionárias, enquanto os bilionários homens 75% são self-made. A correlação entre gênero e self-made, é positiva e média entre as variáveis, visto que o coeficiente de contingência é 0,44.

A distribuição de herdeiros e self-made por indústria é desigual. Apesar de todas elas possuírem maioria em self-made, com exceção o setor “Diversified”, com 44,39% de self-made contra a indústria de tecnologia e telecomunicações que apresentam 92,99% e 83,87% de bilionários self-made respectivamente. Além disso, certos setores apresentam uma certa concentração de patrimônio apesar de conter menos bilionários, como, o setor da moda que se encontra no segundo setor com maior patrimônio de bilionários tem 106 bilionários há menos do que o setor financeiro com 372 bilionários e um patrimônio 5,52% menor que o da moda.

Analisando o gráfico de dispersão entre a quantidade de bilionários, população e o PIB percebe-se que há uma relação linear entre eles. Com a o modelo de regressão linear entre essas variáveis o PIB é explicado em 0,94 pelas variáveis observadas quantidade de bilionários e população, no qual o país que não tem bilionários e nem população tem um PIB esperado de U\$113,1 bilhões e aumentaria em U\$29,9 bilhões para cada bilionários a mais no país e aumentaria em U\$772,7 para cada integrante do país.

Por fim, a distribuição de bilionários entre os países demonstra um valor esperado no qual existe uma concentração nas duas maiores economias do mundo, Estados Unidos e China, no qual os dois juntos apresentam 48,37% dos bilionários dos 78 países na amostra. Além disso, o patrimônio dos bilionários desses dois países é 9,52% maior do que o patrimônio dos bilionários de todos os outros países.

Referências

Base de dados - <https://www.kaggle.com/datasets/nelgiryewithana/billionaires-statistics-dataset>

Participação dos integrantes

Matheus Donito (22/1018307) - Relatório, gráficos e formatação

João Pedro Vieira (22/1018281) - Relatório e tabelas

Lara Siqueira (22/1029122) - Apoio