

Análise dos concursos da Mega-Sena

Universidade Federal Rural de Pernambuco
Computação para Análise de Dados – 2018.1

Bruno Henrique Pereira Marques
Danielly de Moura Borba Queiroz

Recife, 12 de Agosto de 2018

1. Introdução

No Brasil, milhões de pessoas jogam nas loterias federais, seja com baixa frequência ou tem o costume de apostar. Assediadas pelos altos valores das estimativas do prêmio, ainda sim, sabem que a probabilidade de acertar as seis dezenas da Mega-Sena é de 1 em 50.063.860 (cinquenta milhões, sessenta e três mil e oitocentos e sessenta), apostando em apenas um jogo com 6 dezenas que custa R\$3,50. É possível reduzir esta proporção para 1 em 10.003 (dez mil e três) porém, esta aposta custará R\$17.517,50 [Ref. 1].

O valor arrecadado em cada concurso da Mega-Sena é dividido nas seguintes partes: 54,7% para Programas Sociais do Governo e 45,3% para a Premiação. Ainda há um desconto de 13,8% da parte da premiação para o imposto de renda. Por fim, o resultado desta operação é dividido em: 35% para 6 acertos, 19% para 5 acertos, 19% para 4 acertos, 22% acumulado para o próximo concurso e 5% acumulado para o concurso no final do ano. Apesar da porcentagem final do prêmio ser baixa, o número de apostadores tem aumentado [Ref. 2].

Através dos dados históricos de todos os concursos da maior loteria do Brasil [Ref. 3], realizados desde o primeiro (em 11/03/1996) até o último concurso (número 2000) do ano de 2017, esperamos revelar dados como: crescimento do número de ganhadores, crescimento do valor dos prêmios, estado e região com maior número de ganhadores, estado e região com maior valor premiado, e relações entre a ocorrência das dezenas.

Para explorar os dados da Mega-Sena, será utilizado o arquivo no formato HTM disponibilizado no site oficial da loteria [Ref. 1]. Esta é a única forma que os dados são encontrados. A lista a seguir apresenta as principais variáveis contidas no arquivo.

- Número do concurso
- Data do concurso
- Primeira dezena sorteada
- Segunda dezena sorteada
- Terceira dezena sorteada
- Quarta dezena sorteada
- Quinta dezena sorteada
- Sexta dezena sorteada
- Ganhadores da sena
- Cidade dos ganhadores
- Estado dos ganhadores
- Valor de rateio da sena
- Valor acumulado

A proposta deste trabalho é apresentar gráficos de interesse, representados por ano, por estado, por região, pelos valores dos prêmios e com base nas dezenas sorteadas. Esses gráficos ajudarão a apresentar o crescimento da loteria no decorrer dos anos, como o cenário está montado de acordo com os ganhadores no Brasil e como a ocorrência das dezenas está distribuída. Além disso, será apresentada previsões de estimativas de prêmio, valores acumulados e estimativa para a Mega-Sena da virada.

2. Pacotes requeridos

A tabela seguinte apresenta os pacotes que são necessários para executar o código sem erros.

Nome do Pacote	Descrição
Tidyr	Pacote para organização de dados.
Dplyr	Pacote para organização de dados.
Maps	Desenha linhas e polígonos conforme um banco de dados de mapas.
Maptools	Leitura de ESRI <i>shapefiles</i> de mapas.
Mapproj	Gerencia mapas, projeções e <i>grids</i> .
Ggmap	Gerencia mapas baseados em ggplot2, com suporte a Google Maps.
Rgdal	Conjunto de operações para projeções.
DMwR	Biblioteca com o KNN.

3. Preparação dos dados

Os dados inicialmente foram baixados pelo link [Ref.1] onde se encontram no formato HTML. Cada variável está localizada e uma coluna do código HTML conforme a figura 1.

Concurso	Data_Sorteio	1ª	2ª	3ª	4ª	5ª	6ª	Arrecadação_Total	Ganhadores_Sena	Cidade	UF	Rateio_Sena	Ganhadores_Quina	Rateio_Quina	Ganhadores_Quadra	Rateio_Quadra	Acumulado	Valor_Acumulado
1	11/03/1996	04	05	30	33	41	52	0,00	0			0,00	17	39.158,92	2016	330,21	SIM	1.714.650,23
2	18/03/1996	09	37	39	41	43	49	0,00	1		PR	2.307.162,23	65	14.424,02	4488	208,91	NÃO	0,00
3	25/03/1996	10	11	29	30	36	47	0,00	2		RN SP	391.192,51	62	10.515,93	4261	153,01	NÃO	0,00
4	01/04/1996	01	05	06	27	42	59	0,00	0			0,00	39	15.322,24	3311	180,48	SIM	717.080,75
5	08/04/1996	01	02	06	16	19	46	0,00	0			0,00	98	5.318,10	5399	96,53	SIM	1.342.488,85
6	15/04/1996	07	13	19	22	40	47	0,00	0			0,00	109	7.214,66	7147	110,03	SIM	2.286.166,33
7	22/04/1996	03	05	20	21	38	56	0,00	0			0,00	100	8.746,05	5736	152,48	SIM	3.335.692,28
8	29/04/1996	04	17	37	38	47	53	0,00	0			0,00	86	16.084,11	5262	183,40	SIM	4.493.746,19
9	06/05/1996	08	43	54	55	56	60	0,00	0			0,00	17	60.043,79	2175	469,31	SIM	5.719.041,49
10	13/05/1996	04	18	21	25	38	57	0,00	0			0,00	251	6.636,45	12590	132,35	SIM	13.334.769,81
11	20/05/1996	15	25	37	38	58	59	0,00	1		RJ	15.591.365,07	148	12.706,05	9442	199,16	NÃO	0,00
12	27/05/1996	04	16	19	20	27	43	0,00	0			0,00	108	6.791,17	6454	113,64	SIM	880.136,24
13	03/06/1996	18	32	47	50	54	56	0,00	0			0,00	43	20.543,98	2934	301,09	SIM	1.940.205,35
14	10/06/1996	02	16	23	27	47	53	0,00	0			0,00	85	10.458,19	5323	167,00	SIM	3.006.941,32
15	17/06/1996	12	33	35	51	52	60	0,00	0			0,00	57	17.919,89	3931	259,84	SIM	4.232.067,35
16	24/06/1996	20	32	34	49	58	60	0,00	0			0,00	31	33.307,13	3250	317,70	SIM	5.471.696,62
17	01/07/1996	06	10	13	19	20	51	0,00	1		SP	6.789.869,08	144	7.628,37	9376	117,16	NÃO	0,00
18	08/07/1996	23	27	36	37	42	56	0,00	0			0,00	52	13.763,52	4484	159,61	SIM	858.843,48
19	15/07/1996	05	10	12	24	25	60	0,00	0			0,00	159	5.380,14	9250	92,48	SIM	10.506.266,73
20	22/07/1996	11	25	28	30	33	51	0,00	0			0,00	72	15.486,97	5726	194,67	SIM	11.944.540,63
21	29/07/1996	06	33	36	40	49	53	0,00	0			0,00	63	17.694,19	5030	221,62	SIM	13.182.021,44
22	05/08/1996	01	09	31	38	46	56	0,00	0			0,00	47	27.212,90	4965	262,90	SIM	14.716.828,84
23	12/08/1996	17	37	39	51	52	59	0,00	0			0,00	74	20.198,46	4897	305,23	SIM	16.510.452,00
24	19/08/1996	01	08	14	28	33	43	0,00	1			18.661.679,61	227	7.897,31	13486	132,93	NÃO	0,00
25	25/08/1996	24	43	50	54	55	56	0,00	0			0,00	28	26.338,50	2741	269,06	SIM	884.973,50

Figura 1: Tabela HTML com os dados da Mega-Sena.

Em seguida, foi realizado uma conversão manualmente do código HTML para CSV utilizando a ferramenta Excel 2013. Dessa forma, passa a ser possível carregar os dados para processamento com a linguagem R. A figura 2 apresenta parte dos dados carregados na ferramenta RStudio. Alguns problemas de conversão são visíveis como a linha 4, por exemplo, a qual está toda preenchida com *NA*, valores de UF separados por vírgula e a formatação da moeda Rateio_Sena.

Concurso	Data_Sorteio	X1_Dezena	X2_Dezena	X3_Dezena	X4_Dezena	X5_Dezena	X6_Dezena	Arrecadação_Total	Ganhadores_Sena	Cidade	UF	Rateio_Sena	Ganhadores_Quina	Ra
1	11/03/1996	4	5	30	33	41	52	0	0	NA	NA	0	NA	17
2	18/03/1996	9	37	39	41	43	49	0	1	NA	PR	2.307.162,23	65	14
3	25/03/1996	10	11	29	30	36	47	0	2	NA	RN,SP	391.192,51	62	10
NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
4	01/04/1996	1	5	6	27	42	59	0	0	NA	NA	0	NA	39
5	08/04/1996	1	2	6	16	19	46	0	0	NA	NA	0	NA	98

Figura 2: Tabela CSV convertida.

Por fim, a remoção, substituição e formatação dos valores foram realizados para que se o manuseio com os dados fique mais coerente. A lista abaixo apresenta a ordem das operações finais da preparação dos dados.

- 3.1. Remoção das linhas com **NA**;
- 3.2. Conversão das datas formatadas como *String* para variáveis numéricas separadas (Dia, Mês e Ano);
- 3.3. Conversão da moeda formatada como *String* para valor numérico decimal;
- 3.4. Conversão dos valores SIM e NAO (*String*) para TRUE e FALSE (*Logical*);
- 3.5. Agrupamento das variáveis X_n Dezenas (n é a ordem a qual a dezena foi sorteada) para um *dataframe* organizado com número do concurso, ordem do sorteio e dezena – figura 3;
- 3.6. Separação dos valores da variável UF para um *dataframe* organizado com número do concurso e UF – figura 4;
- 3.7. Criação de um *dataframe* organizado para as regiões do Brasil – figura 5;
- 3.8. Persistência dos *dataframes* em quatro arquivos CSV.

Concurso	Ordem_Sorteio	Dezena
1	1	4
2	1	9
3	1	10
4	1	1
5	1	1
6	1	7

Figura 3: clean-data-dezenas.csv

Concurso	UF
2	PR
3	RN
3	SP
11	RJ
17	SP
29	RS

Figura 4: clean-data-estados.csv

Regioes	Regiao_id	UF
NO	3	AM
NO	3	RR
NO	3	AP
NO	3	PA
NO	3	TO
NO	3	RO

Figura 5: clean-data-regioes.csv

4. Análise exploratória dos dados

Com base no *dataframe* que contém os dados gerais da Mega-Sena, podemos observar no gráfico da figura 6, o crescimento do valor do prêmio pago por ano em milhões de reais. Em 1996 o valor total que a loteria pagou como premiação foi cerca de 81 milhões de reais. Apesar da queda em 2016, o gráfico se mostra crescente chegando a pagar cerca de 1,3 bilhão de reais em 2015 – ano que mais premiou na loteria. Vale ressaltar que o valor do prêmio está relacionado ao número de apostas, sendo assim, a popularidade da Mega-Sena continua crescendo.

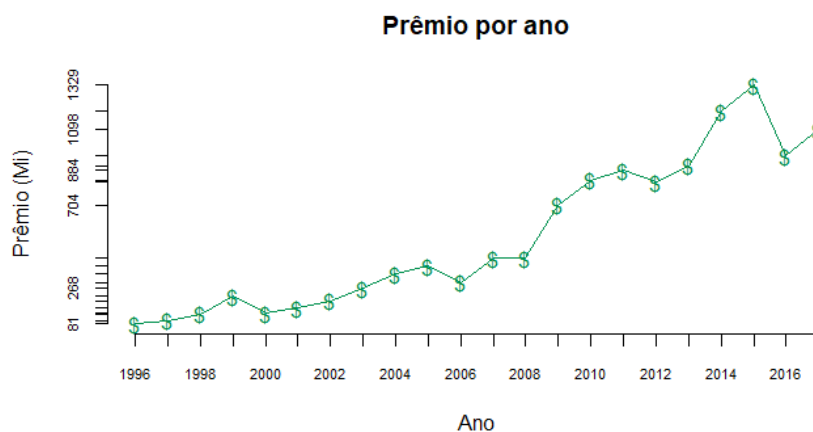


Figura 6: Prêmio por ano

Sabemos que a probabilidade de ganhar na Mega-Sena é baixíssima, porém de acordo com o contexto histórico, a figura 7 apresenta um número crescente de ganhadores. No gráfico anterior, vimos que em 2015 foi o ano com maior premiação, porém, o ano que houve mais ganhadores foi em 2012 com 54 pessoas. Isto também significa que em 2015 o valor do rateio (divisão do prêmio pelo número de ganhadores) foi um dos maiores.



Figura 7: Ganhadores por ano

Diante de problemas atrelados à corrupção no Brasil, a confiança da população quando se envolve dinheiro em diversos setores é limitada. Na Mega-Sena existem pessoas que acreditam que os concursos são fraudados [Ref. 4]. Por ser um jogo aleatório, todas as 60 dezenas devem possuir a mesma probabilidade de serem sorteadas, resultando em uma distribuição uniforme. Para verificar parte da confiabilidade da loteria, a figura 8 apresenta o gráfico com a probabilidade da ocorrência das dezenas. Note que a linha de densidade se mantém uniforme para todas as dezenas, ou seja, a probabilidade de sorteio se mantém bem próxima para qualquer dezena. Ainda assim, existem dezenas as quais foram mais e menos sorteadas em quantidade exata. As seis mais sorteadas são: 5, 53, 10, 23, 4 e 51 respectivamente. As menos sorteadas são: 26, 55, 22, 21, 9, 25 respectivamente.

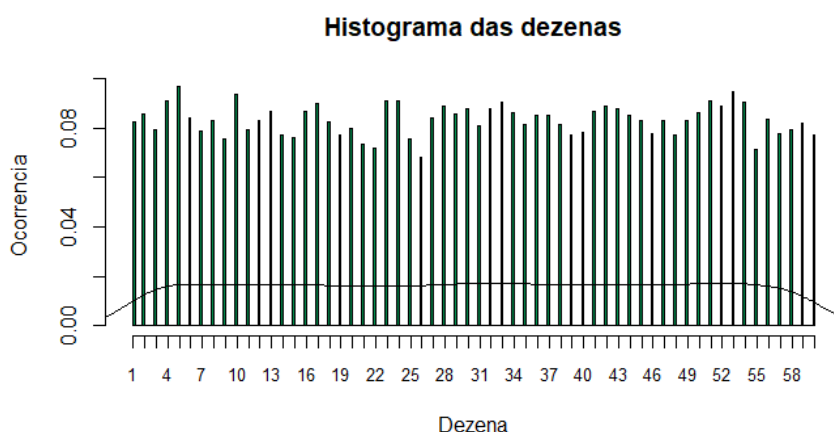


Figura 8: Histograma das dezenas

Alguns entusiastas e fanáticos pela loteria costumam elaborar fórmulas para tentar calcular o jogo que pode mudar a vida do apostador [Ref. 5]. A partir dos dados é possível verificar, na figura 9, a ocorrência das dezenas as quais foram sorteadas em concurso com

ganhadores. Também é possível verificar quais foram as dezenas mais e menos apostadas entre os ganhadores. As seis mais apostadas são: 5, 10, 33, 3, 1 e 17 respectivamente. As menos apostadas são: 60, 31, 26, 55, 50, e 30 respectivamente.

Histograma das dezenas premiadas

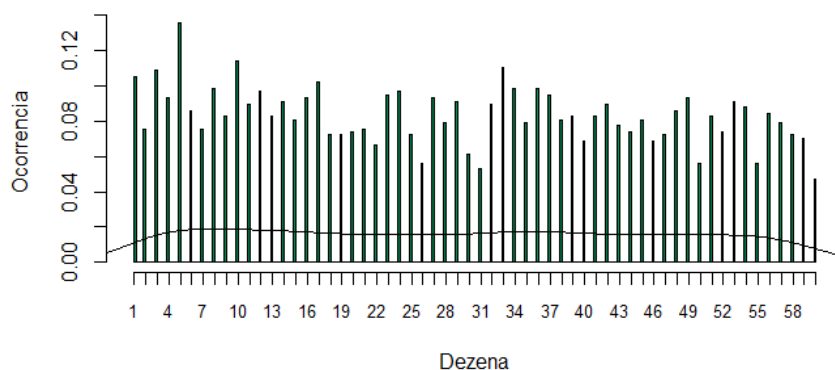


Figura 9: Histograma das dezenas premiadas

De todas as formas, seja apostando números aleatórios, apostando números combinados ou qualquer outra combinação de dezenas, apostadores da Mega-Sena possuem a esperança de ganhar. Para os pessimistas, o gráfico da figura 10 mostra que de 2000 concursos, 24% houveram ganhadores. O número de ganhadores em todos os sorteios da Mega-Sena foi 685 pessoas até o final de 2017. Também podemos observar como está o cenário dos premiados no Brasil.

Ganhadores x Acumulado

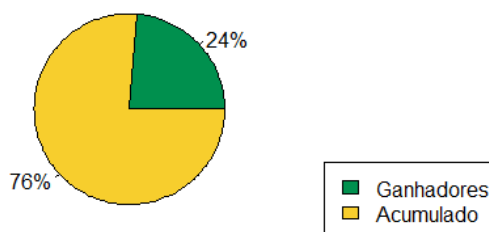


Figura 10: Ganhadores x Acumulado

A figura 11 apresenta o número e a projeção de ganhadores por estados. Note que São Paulo é o estado com maior número de vencedores. Isso pode estar relacionado a cultura de apostar com frequência, como também pode estar associado ao número de lotéricas bem como a divulgação da loteria. Infelizmente os dados disponibilizados não contém a quantidade de apostas.

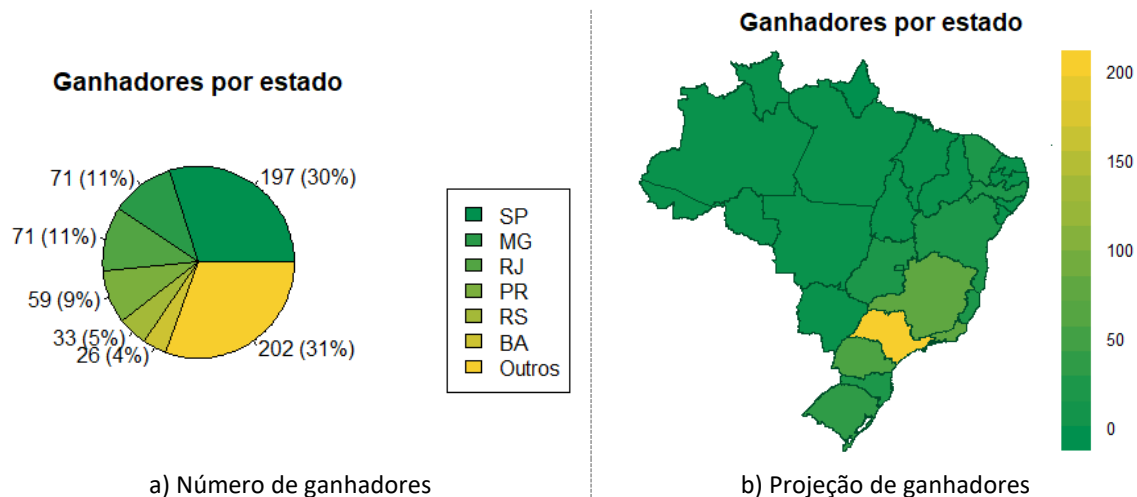


Figura 11: Ganhadores por estado

A figura 12 apresenta o número e a projeção de ganhadores por região. Note que a região Sudeste possui maior número de vencedores.

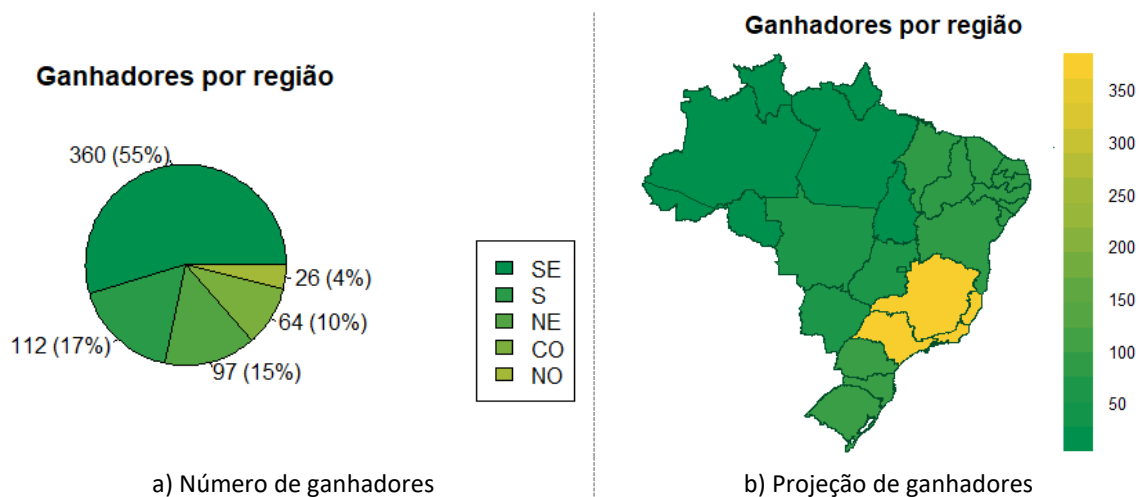


Figura 12: Ganhadores por região

Além disso, realizamos uma análise exploratória dos dados, através da regressão linear, com o objetivo de analisar os dados estaticamente, ou seja, verificar a existência de uma relação entre uma variável dependente com outras variáveis independentes. Para realizar a regressão linear para tentar prever executamos os seguintes passos:

- I. Verificar correlação entre as variáveis;
- II. Encontrar a Equação do modelo Linear;
- III. Obter gráfico de regressão;
- IV. Avaliar se o modelo é estaticamente significativo;
- V. Analisar os Pressupostos;
- VI. Executar as Predições.

Nosso primeiro teste foi realizado para estimar os valores de prêmios da loteria. Utilizando o *dataframe* que contém os dados gerais da Mega-Sena, e a execução da regressão linear, obtemos esse resultado:

Primeiro que a correlação entre as variáveis Estimativa_Premio e Concurso possuem um grau fraco de associação, porque o valor da correlação foi em torno de 0.45. Também obtemos o gráfico de dispersão com a linha de regressão, onde cada ponto representam as estimativas dos valores dos prêmios em um determinado concurso, como mostra a figura 13.

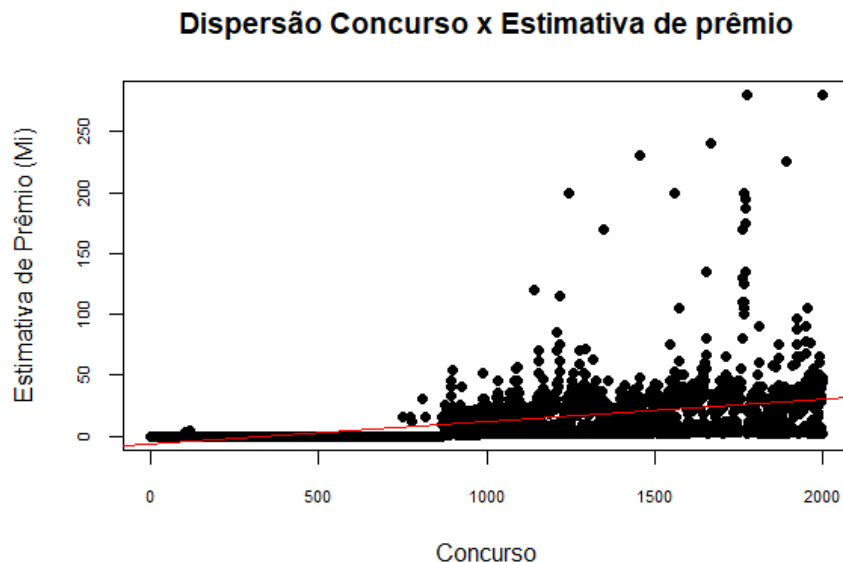


Figura 13: Regressão Linear Concurso x Estimativa Prêmio

Avaliamos o nosso modelo para verificar se variável X (Concurso) pode ou não ser utilizada para prever Y (Estimativa_Premio). Então, se o valor de p associado ao coeficiente for igual ou menor que o nível de significância adotado, então aceitamos que existe uma relação estatísticas entre as variáveis. Como o $\Pr(>|t|) = 2.22e-16$, que extraímos do nosso modelo é menor do 0.05, podemos dizer então que existe uma relação estatísticas. E como $R = 0.2031$ o modelo de regressão não explicou grande parte da variabilidade no resultado.

Além disso foi analisado os pressupostos, verificamos se as variáveis são homogêneas e a normalidade dos dados, onde foi possível observar que como $p\text{-value} = 2.220446049e-16 < 0.05$, podemos concluir que os resíduos não são provenientes de uma distribuição normal. Dessa forma, testamos as previsões para os concursos 2001, 2020, 2068 e com o isso podemos dizer que para o concurso 2001 o valor estimado para o prêmio será: 30449279.0395198. Em 2020 o valor estimado para o prêmio: 30802203.8788510 e em 2068 o valor estimado para o prêmio: 31693803.4729509.

O segundo teste realizado foi para estimar o Valor Acumulado em concursos futuros, realizamos os mesmos passos do teste anterior, só que agora utilizando as variáveis Valor_Acumulado e Concurso do nosso *dataframe* que contém os dados gerais da Mega-Sena. Com a realização do teste de correlação entre essas variáveis, podemos dizer que a correlação foi 0.28, então elas possuem um grau fraco de associação. Assim como o outro teste, também obtemos o gráfico de dispersão com a linha de regressão, onde cada ponto representam os valores acumulados em um determinado concurso, a figura 14 mostra isso.

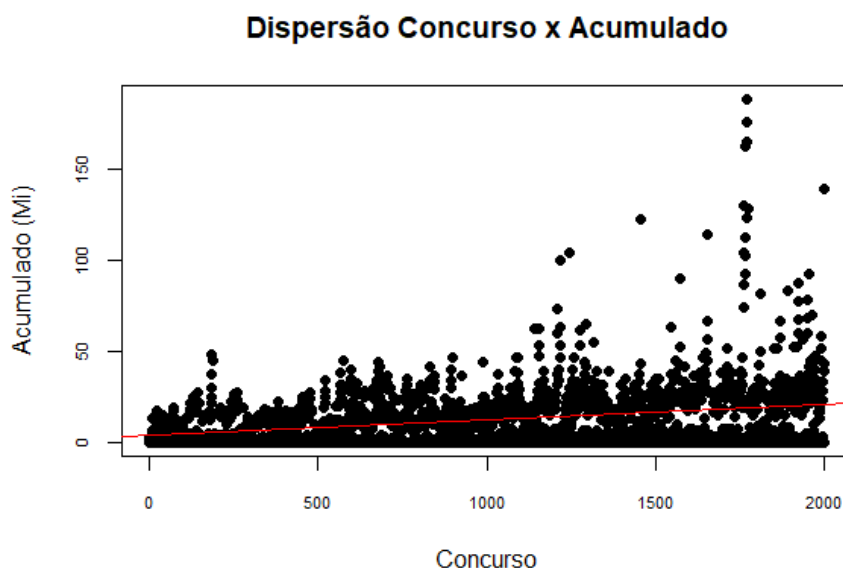


Figura 14: Regressão Linear Concurso x Valor Acumulado

Avaliamos o nosso modelo, como o valor de $\Pr(>|t|) = 2.22e-16$ que extraímos do nosso modelo é menor do 0.05, podemos dizer então que existe uma relação estatísticas. E como $R = 0.077$ o modelo de regressão não explicou grande parte da variabilidade no resultado. Analisamos também os pressupostos através das verificações da homogeneidade e normalidade dos dados, com resultado dessa análise como $p\text{-value} < 2.2e-16$, os dados não seguem uma distribuição normal. Além disso foi analisado os pressupostos, verificamos se as variáveis são homogêneas e a normalidade dos dados, onde foi possível observar que como $p\text{-value} = 2.220446049e-16 < 0.05$, podemos concluir que os resíduos não são provenientes de uma distribuição normal. Dessa forma, testamos as previsões para os concursos 2001, 2020 e 2068 e com o isso podemos dizer que para o concurso 2001 o valor acumulado está estimado em: 20455133.6680038. Em 2020 o valor acumulado está estimado em: 20612926.8900945 e em 2068 valor acumulado está estimado em: 21011562.3985342.

E por fim dos testes das previsões utilizando regressão linear, queremos estimar o Valor Acumulado Mega da Virada. E como primeiro passado observamos a correlação entre as variáveis Acumulado_Mega_da_Virada e Concurso, como o valor da correlação foi 0.70 então ela possui um grau moderado ou bom de associação. Logo abaixo, a figura 15 apresenta onde cada ponto que representam as estimativas dos valores acumulados da mega da virada em um determinado concurso. A partir disso, avaliamos o nosso modelo, como o valor de $\Pr(>|t|) = 2.22e-16$ é menor do 0.05, podemos dizer então que existe uma relação estatísticas. E como $R = 0.49$, podemos dizer o modelo de regressão não explicou grande parte da variabilidade no resultado. Analisamos também os pressupostos através das verificações da homogeneidade e normalidade dos dados.

Até que podemos testar as previsões para os concursos 2001, 2020 e 2068 e com o isso podemos dizer que para o concurso 2001 o valor acumulado da Mega da Virada está estimado em: 42224816.6541231. Em 2020 o valor acumulado da Mega da Virada está estimado em: 42728663.7842579 e em 2068 o valor acumulado da Mega da Virada em: 44001540.7445984.

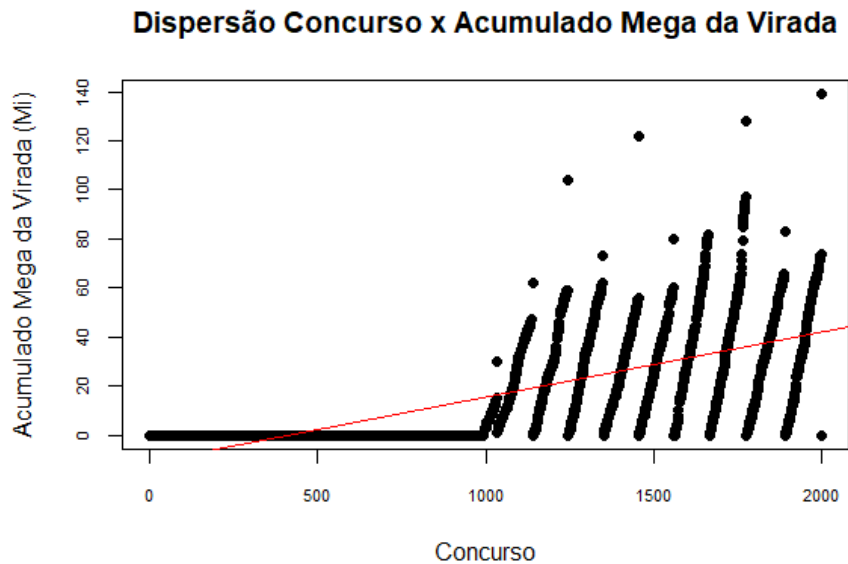


Figura 15: Regressão Linear Concurso x Acumulado Mega da Virada

Nos dias de hoje é imprescindível o uso de algoritmos de inteligência artificial para classificação de objetos como também, a previsão de eventos. Neste trabalho foi utilizado o algoritmo *K-Nearest Neighbors* para a previsão de qual dia um estado irá ganhar. Para isso, foi utilizado o dia e o mês como características de treinamento. Dessa forma, se quisermos saber qual estado ganharia em um determinado dia, seria preciso informar apenas o dia e o mês que se deseja descobrir. A Figura 16 apresenta a previsão dos estados ganhadores para todos os dias do mês de março. Note que a saída do algoritmo prevê que o vencedor dos concursos nos dias 16 e 21 do mês de março, serão de São Paulo. Da mesma forma, para o dia 28 de março, o algoritmo prevê que o vencedor será do Paraná.



Figura 16: Previsão do KNN para estados ganhadores

Outra versão implementada do *KNN* neste trabalho foi para prever se uma aposta com 6 dezenas seria capaz de ganhar na Mega-Sena, bem como sugerir jogos. A figura 17 apresenta as entradas e saída do algoritmo. A simulação mostra uma aposta com as dezenas 5, 6, 14, 22, 39 e 56 e o resultado do algoritmo afirma que este jogo não acumularia o concurso, ou seja, ele tem grande chance de acertar a Mega-Sena.

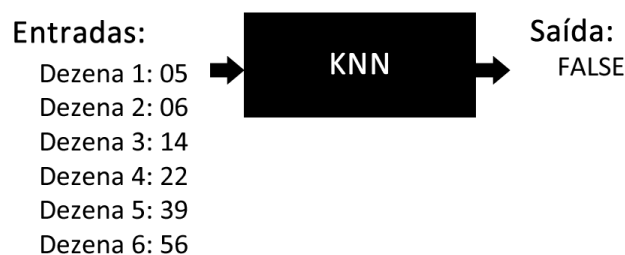


Figura 14: Previsão do KNN para estados ganhadores

5. Conclusões

5.1. Problema:

A análise deste projeto é realizada para que os clientes da Mega-Sena possam avaliar o sistema apresentando gráficos de interesse. Esses gráficos ajudam a apresentar o crescimento da loteria no decorrer dos anos, como o cenário está montado de acordo com os ganhadores no Brasil e como a ocorrência das dezenas está distribuída. Podendo também realizar previsões de estimativas de prêmio, valores acumulados e estimativa para a Mega-Sena da virada. Além disso tudo, este projeto utiliza inteligência artificial para prever quando um estado ganharia na loteria e para sugerir possíveis jogos premiados.

5.2. Metodologia:

Com objetivo de entender melhor o problema, primeiro aplicamos um gráfico para avaliar o crescimento do valor do prêmio pago por ano em milhões de reais. Em seguida, mesmo com o conhecimento que a probabilidade de ganhar na Mega-Sena é baixíssima, construímos um gráfico para verificar como vem se comportando o número de ganhadores ao decorrer dos anos. Além disso, para verificar parte da confiabilidade da loteria construímos um gráfico com as ocorrências das dezenas nos concursos. Em seguida, construímos um gráfico para verificar a porcentagem dos concursos que tiveram ganhadores e os que foram acumulados. Também, utilizamos as informações sobre os concursos e os estados ganhadores da base, para analisar qual estado que mais ganhou e também realizamos o mesmo procedimento só que para regiões, verificando qual região que obteve o maior número de ganhadores. A partir disso, nos aprofundamos nossa análise realizando testes estatísticos para prever possíveis valores de prêmios futuros, valores acumulados e valores acumulados na Mega-Sena da virada. Por fim, estendemos o projeto utilizando o algoritmo *KNN* para aprendizagem de máquina de modo que se possa fazer previsões de quando um estado irá ganhar na loteria, bem como sugerir jogos com base no histórico.

5.3. Insights:

- a) O valor do prêmio pago por ano em milhões de reais vem crescendo;
- b) O número de ganhadores também está crescendo;
- c) Apesar da desconfiança de fraude em relação aos jogos, foi mostrado através do gráfico que probabilidade de sorteio se mantém bem próxima para qualquer dezena.
- d) As dezenas que foram mais apostadas entre os ganhadores são respectivamente: 5, 10, 33, 3, 1 e 17. As menos apostadas são 60, 31, 26, 55, 50 e 30, respectivamente.
- e) 24% dos concurso houveram ganhadores, para os entusiastas isso pode ser um número relevante visto que a probabilidade de ganhar na loteria é baixíssima.
- f) O estado de São Paulo tem o maior número de vencedores. Isso pode estar relacionado a cultura de apostar com frequência.
- g) Nos jogos da Mega-Sena, a região Sudeste possui maior número de vencedores.
- h) Estimamos que o valor do prêmio nos próximos concursos pode chegar a cerca de 31 milhões de reais.
- i) O valor acumulado estimado para os próximos concursos pode ser de cerca de 21 milhões de reais.

- j) O valor acumulado para a Mega da Virada pode chegar a 44 milhões de reais.

5.4. Implicações:

Esta análise pode ser utilizada para compreender como esse jogo tão popular no nosso país, como é a Mega-Sena, se comporta. Com resultados das análises, podemos observar que o prêmio vem crescendo a cada ano, isto devido ao aumento das apostas. Além disso, observamos que houve um crescimento do número de ganhadores com o decorrer dos anos, mesmo que a probabilidade de ganhar na Mega-Sena seja baixíssima. Apesar deste crescimento, observamos que apenas 24% dos concursos houveram ganhadores e o resto foi acumulado.

Em relação ao sorteio das dezenas, podemos observar que a probabilidade se mantém bem próxima para qualquer dezena, ajudando na confiabilidade da Mega-Sena. A existência das diferenças entre as dezenas se dá pelo fato de que a loteria tem base em sorteios aleatórios.

Podemos verificar também que devido a quantidade altas de apostas, o estado de São Paulo foi o que obteve maior número de ganhadores e em relações as regiões do Brasil, a que teve o maior número de ganhadores foi a Região Sudeste.

Com todo o crescimento de aposta, prêmio, ganhadores que podemos notar através de previsões que o valor do primeiro, o valor acumulado e o valor acumulado da Mega-Sena tende aumentar.

5.5. Limitações

Este projeto de análise foi limitado devido à falta de informações relevantes para cálculos mais embasados, como não ter informações da quantidade de apostas por estados. Futuramente poderíamos buscar fontes que possam disponibilizar dados sobre a Mega-Sena, para que seja possível incluir dados como quantidade de apostas realizadas em todo Brasil.

6. Referências

Ref. 1 – Caixa Econômica Federal – Mega-Sena.

<http://loterias.caixa.gov.br/wps/portal/loterias/landing/megasena/>

Ref. 2 – G1 – Número de ganhadores da Mega da Virada foi recorde; veja estatísticas.

<https://g1.globo.com/economia/noticia/numero-de-ganhadores-da-mega-da-virada-foi-recorde-veja-estatisticas.ghtml>

Ref. 3 – IQ Sorteios – Conheça as três loterias mais populares do Brasil.

<https://www.iq360.com.br/sorteios/artigos/as-loterias-mais-populares-do-brasil>

Ref. 4 – UOL Notícias – Coincidência inédita na Mega-Sena provoca dúvidas nas redes sociais; matemáticos explicam.

<https://noticias.uol.com.br/cotidiano/ultimas-noticias/2018/06/25/improvavel-impossivel-resultado-diferente-incentiva-teorias-de-fraude-na-mega.htm>

Ref. 5 – UOL Notícias – O que dizem os matemáticos sobre as ditas fórmulas certeiras e dicas infalíveis para se ganhar na loteria.

<https://noticias.uol.com.br/ciencia/ultimas-noticias/bbc/2017/12/29/o-que-dizem-os-matematicos-sobre-as-ditas-formulas-certeiras-e-dicas-infaliveis-para-se-ganhar-na-loteria.htm>