

ENSINANDO E APRENDENDO

UNIFOR - Universidade de Fortaleza

MBA EM CIÊCIA DE DADOS

JOSÉ VALCLEMIR RODRIGUES DA SILVA

mortes_por_covid19 20 -0mar 23 abr 06 abr 13 mar 16 mar 30 dias In [137]: | #Filtrando por datas data_inicial <- as.Date('2020-05-01')</pre> data_final <- as.Date('2020-05-25')</pre> dim(subset(df_caso_full,df_caso_full\$date >= data_inicial & df_caso_full\$date <= data_final))</pre> 0 16 In [138]: | sum_obtos_gerais_por_covid <- setNames(</pre> aggregate(x=obitos_gerais\$new_deaths, by=list(Category=obitos_gerais\$state),

FUN=sum), c('uf', 'quantidade')) # Basic line plot with points options(repr.plot.width = 10, repr.plot.height = 8) ggplot(data=sum_obtos_gerais_por_covid, aes(x=uf, y=quantidade, group=1)) + geom_line()+ geom_point() 150 100-

In [139]: quantidade 50-BA CE DF ES GO MA MG MS MT PA PB PE PI PR RJ RN RO RR RS SC In [140]: | df_obito_cartorio <- read.csv(file = 'obito_cartorio.csv')</pre> #Converte factor to date newDates <- as.Date(df obito cartorio\$date, format= "%Y-%m-%d") df_obito_cartorio <- mutate(df_obito_cartorio, date= newDates)</pre> df_obito_cartorio['ano_mes'] <- format(as.Date(df_obito_cartorio\$date, format="%d/%m/%Y"),"%Y-%m")</pre> In [141]: df_obito_cartorio

state new_deaths_pneumonia_2019 new_deaths_pneumonia_2020 new_deaths_respiratory_failure_2019 new_deaths_respiratory_failu date 2020-AC 3 3 01-01 2020-AL7 0 12 01-01 2020-AM2 0 0 01-01 2020-ΑP 0 0 0 01-01 2020-BA 10 0 31 01-01 2020-CE 12 0 9 01-01 2020-DF 10 0 2 01-01 2020-ES 17 0 15 01-01 2020-GO 16 0 14 01-01 2020-MA 6 0 10 01-01 2020-

MG 51 0 47 01-01 2020-MS 6 0 5 01-01 2020-MT 5 0 3 01-01 2020-PA 3 0 10 01-01 2020-PB 9 0 10 01-01 2020-PΕ 17 0 17 01-01 2020-Ы 7 0 7 01-01 2020-PR 24 0 13 01-01 2020-RJ67 0 39 01-01 2020-RN11 0 3 01-01 2020-RO 3 0 3 01-01 2020-RR 1 0 0 01-01 2020-RS 28 0 29 01-01 2020-SC 16 0 18 01-01 2020-SE 8 0 8 01-01 2020-SP 199 0 114 01-01 2020-TO 1 0 4 01-01 2020- AC 0 2 0 01-02 2020-AL01-02 2020- AM 3 5 3 01-02

---2020-7 0 7 SE 04-16 2020-SP 202 4 120 04-16 2020-0 3 TO 1 04-16 2020- AC 0 1 04-17 2020-4 0 8 ΑL 04-17

2020-7 0 AM 3 04-17 2020-ΑP 1 0 0 04-17 2020-ВА 20 0 31 04-17 2020-CE 0 27 26 04-17 2020-DF 0 6 04-17 2020-17 0 9 ES 04-17 2020-GO 16 0 12 04-17 2020-0 5 MA 14 04-17 2020- MG 54 0 38 04-17 2020-0 1 MS 14 04-17 2020- MT 5 0 04-17

2020-5 0 7 PA 04-17 2020-РΒ 14 0 15 04-17 2020-0 PΕ 23 14 04-17 2020-Ы 0 11 04-17 2020-0 PR 20 31 04-17 2020-RJ0 68 42 04-17 2020-0 10 6 RN04-17 2020-RO 0 0

04-17 2020-0 0 RR 1 04-17 2020-RS 0 29 21 04-17 2020-

0

0

0

0

sum_mortes_por_estado <- df_obito_cartorio[, c('state', 'deaths_covid19')][df_obito_cartorio[['deaths_</pre>

In [143]: | df1 = df obito cartorio[, c("date", "deaths covid19")]#[which(df obito cartorio\$state == 'CE'),] df2 = df_caso_full[, c("date", "new_deaths")]#[which(df_caso_full\$state == 'CE'),]

In [144]: sum_acumulado_obitos_cartorio <- setnames(aggregate(df1\$deaths_covid19,</pre>

long <- reshape2::melt(df_final, id.vars = "date")</pre>

long <- setnames(long, c("date", "obitos", "value"))</pre>

scale_y_log10(oob = scales::squish_infinite) +

x = "DATA", y = "OBITOS ACUMULADOS")

In [145]: $plot \leftarrow ggplot(long, aes(x = (date), y = value,$

sum_acumulado_obitos_oficiais <- setnames(aggregate(df2\$new_deaths,</pre>

group = obitos, colour = obitos)) +

"Transformation introduced infinite values in continuous y-axis"

Comparação entre os óbitos oficiais e registrados em cartório

labs(title = "Comparação entre os óbitos oficiais e registrados em cartório",

#df obito cartorio[, c("date", "deaths covid19")][which(df obito cartorio\$state == 'CE'),]

df_final = merge(sum_acumulado_obitos_cartorio, sum_acumulado_obitos_oficiais, by = "date")

by=list(Category=df1\$date), FUN=sum), c("date", "cartorio"))

by=list(Category=df2\$date), FUN=sum), c("date", "oficiais"))

plot + scale_x_date(date_labels = "%d-%m") + theme(axis.text.x = element_text(angle = 45, hjust = 1))

DATA

"new_deaths_respiratory_failure_2020",

"new_deaths_covid19")]), 1)

"last available confirmed per 100k inhabitants",

sum_acumulado_100k_hab <- (aggregate(df_100k_hab\$last_available_confirmed_per_100k_inhabitants,</pre>

plot + scale x date(date labels = "%d-%m") + theme(axis.text.x = element text(angle = 45, hjust = 1))

DATA

df_muni_geo_ref <- df_municipios[df_municipios\$codigo_ibge %in% c(df_caso_full\$city_ibge_code[df caso</pre>

Obs: Ao exportar para PDF, vai da problema para visualizar, pois a library que estou utilizando, utiliza javascript. Enviarei o pdf e

FUN=sum))

plot <- $ggplot(sum_acumulado_100k_hab, aes(x = (date), y = x,$ group = uf, colour = uf)) +

subtitle = "Novos óbitos por 100 mil habitantes",

scale_y_log10(oob = scales::squish_infinite) +

x = "DATA", y = "OBITOS ACUMULADOS")

"Removed 136 rows containing missing values (geom_path)."

"state")][which(df_caso_full\$state %in% c('AL', 'BA', 'CE', 'MA', 'PB',

by=list(date=df 100k hab\$date, uf=df 100k hab\$state),

corr <- round(cor(df_obito_cartorio[, c("new_deaths_pneumonia_2020",</pre>

obitos

Corr

1.0 0.5

0.0 -0.5

uf

- AL

ВА - CE MA PB PΕ Ы

SE

cartorio oficiais

14

113

SC

SE

SP

TO

covid19']] == 1,]

FUN=sum)

04-17 2020-

04-17 2020-

04-17 2020-

04-17

Category

AC 11 7 ΑL AM

ВА

DF 2

MS 6 MT 15 PA 2 PB 1

PE 3

PR 3 RJ 1

RN 1 RO 4 RR 9 RS 1

SC 2 SE 4 TO 4

#renomeia colunas

geom_line() +

Warning message:

1000-

OBITOS ACUMULADOS

100-

10-

#Cria matriz de correlação

new_deaths_covid19

new_deaths_respiratory_failure_2020

new_deaths_pneumonia_2020

In [147]: | df 100k hab = df caso full[, c("date",

'PE', 'PI', 'RN', 'SE')),]

geom line() +

Warning message:

100.0

10.0

1.0

0.1

321 6

m1

OBITOS ACUMULADOS

Mortalidade

#theme_ipsum() + theme bw() +

labs(title = "Mortalidade",

Novos óbitos por 100 mil habitantes

In [148]: # Busca todos os casos de covid que levaram ao óbito

os dois mapas no formato html, com os seguintes nomes:

Mapa 2: heat_map_municipios_obitos_por_covid.html

clusterOptions = markerClusterOptions()

#Salva o mapa na pasta local para visualizacao

In [149]: #Plota a clusterização de todos os casos de óbito nos estados

m1 <- leaflet(df muni geo ref) %>% addTiles() %>% addMarkers(

In [150]: #plota mapa de calor de todos os municipios onde houveram obítos por covid

addHeatmap(lng = ~longitude, lat = ~latitude, radius = 8)

saveWidget(m2, 'heat_map_municipios_obitos_por_covid.html', selfcontained = FALSE)

Regressão linear para medir a correlação entre NOVOS CASOS CONFIRMADOS X IDHM

População

estimada -

pessoas

[2019]

11737

14929

62641

54270

17399

11714

Densidade

- hab/km²

[2010]

58.69

98.52

68.31

22.58

6.66

77.71

demográfica

Escolarização

6 a 14

anos -

% [2010]

IDHM

96.7 0.286

96.8 0.349

96.8 0.277

97.2 0.254

97.5 0.256

97.7 0.285

Área

- km²

[2019]

180.833

130.002

842.471

2254.279

2438.563

135.760

In [154]: levels(df muni ce\$city)[levels(df muni ce\$city)=="Santa QuitÃ@ria"] <- "Santa Quitéria" levels(df muni ce\$city)[levels(df muni ce\$city)=="Eusébio"] <- "Eusébio" levels(df muni ce\$city)[levels(df muni ce\$city) == "TianguÃ;"] <- "Tianguá"</pre> levels(df_muni_ce\$city)[levels(df_muni_ce\$city) == "Cariðs"] <- "Cariús"</pre>

levels(df muni ce\$city)[levels(df muni ce\$city) == "Maracanað"] <- "Maracanaú"</pre>

In [153]: $\#df_{muni}ce \leftarrow df_{caso}full[(df_{caso}full['new_deaths'] == 1 & df_{caso}full['state'] == 'CE'),]$

In [156]: d <- d[!rowSums(is.na(d[, c("last available confirmed per 100k inhabitants", "last available death rat

Prefeito Territorial

[2021]

AFONSO

TAVARES

FRANCISCO EDILBERTO

BESERRA BARROSO ANA FLÁVIA

RIBEIRO

ANTÔNIO

ALMEIDA

RAMILSON

ARAUJO

MORAES JOAQUIM

FREIRE

df muni ce <- df caso full[(df caso full['state'] == 'CE'),]</pre>

Seleciona as colunas CIDADE e IDHM e logo em seguida, guarda em numa variável

CARVALHO

NETO

MONTEIRO

LEITE

Mortalidade infantil -

óbitos por

nascidos

mil

vivos [2017]

33.33

14.56

8.81

19.54

10.87

25.42

Receitas

(×1000)

[2015]

28986.22

32881.18

55723.51

93726.44

33372.70

26602.08

- R\$

realizadas

Desp

empenha

- R\$ (×1

[2

2035

2446

8686

7052

2799

2066

m2 <- leaflet(df_muni_geo_ref) %>%

#Salva o mapa na pasta local visualizacao

addTiles() %>%

Carrega planilha de IDH

Visualiza os dados

Abaiara 2300101

Acarape 2300150

Acaraú 2300200

Acopiara 2300309

Aiuaba 2300408

Alcântaras 2300507 alcantarense

Filtra apenas o estado do CEARA

Subistitui caractéres unicode

In [155]: df idh muni <- df idh[, c('city', 'IDHM')]</pre>

Dropa os NA's

e")])),]

d <- merge(df muni ce, df idh muni, by="city")</pre>

head(df idh)

In [152]:

In [151]: | df_idh <- read_excel("idh.xlsx")</pre>

Gentílico [-]

abaiarense

acarapense

acarauense

acopiarense

aiuabense

Mapa 1: mapa_cluste_obitos_nos_estados.html

 $_{\text{full$new_deaths}} = 1]),]$ #Mostra shape do dataset dim(df muni geo ref)

df_municipios <- read.csv("municipios_lat_long.csv")</pre>

Plota mapas de clusterição de casos de obito nos estados

saveWidget(m1, 'mapa_cluste_obitos_nos_estados.html', selfcontained = FALSE)

Assuming "longitude" and "latitude" are longitude and latitude, respectively

#plota a correlação ggcorrplot(corr)

In [146]:

2

In [142]:

17

4

188

by=list(Category=sum_mortes_por_estado\$state),

aggregate(sum mortes por estado\$deaths covid19,

