Scratch

October 11, 2019

```
In [66]: %load_ext rpy2.ipython

         import os
         os.environ['KMP_DUPLICATE_LIB_OK']='True'
```

```
The rpy2.ipython extension is already loaded. To reload it, use:
  %reload_ext rpy2.ipython
```

```
In [67]: import pandas as pd
         import numpy as np

In [68]: mileage = pd.read_csv('./data_reshaped.csv')
         gdp=pd.read_csv('GDP.csv')

In [69]: len(mileage)

Out[69]: 30243

In [70]: m1 = mileage.loc[mileage.area == 'all']
         m1 = m1.groupby(['monthid']).sum().reset_index().sort_values(['monthid'])
         m1 = m1[['monthid','Million_Vehicle_Miles']]
         m1.columns = ['monthid','miles']

In [71]: m1['date_str'] = m1.monthid.map(lambda x: str(x))
         m1['month'] = m1.date_str.str[-2:]
         m1['year'] = m1.date_str.str[:4]

In [72]: m1['rolling'] = m1['miles'].rolling(12).sum()/1000000
         m1= m1.iloc[11:,:].reset_index()
         m1['order'] = m1.index

In [73]: gdp['order'] = gdp.index
         gdp['gpd_shrunk'] = gdp.GDP / 1000

In [74]: m2 = mileage.copy()
         m2['date_str'] = m2.monthid.map(lambda x: str(x))
         m2['month'] = m2.date_str.str[-2:]
         m2['year'] = m2.date_str.str[:4]
```

```
In [75]: m2 = (m2.loc[m2.year == '2018'].groupby(['State','area'])
                .agg({'Million_Vehicle_Miles':['sum']})
               )

In [76]: m2 = m2.reset_index()

In [77]: m2_urb = m2.loc[m2.area == 'urban']
         m2_rur = m2.loc[m2.area == 'rural']
         m2_rur = m2_rur.reset_index()
         m2_urb = m2_urb.reset_index()

In [78]: ur_ratios= (m2_urb[('Million_Vehicle_Miles','sum')] /
         (m2_rur[('Million_Vehicle_Miles','sum')] + 0.01))

In [79]: ratios_2018 = pd.DataFrame(ur_ratios).join(m2_urb.State)
         ratios_2018.columns = ['ratio','State']
```

```
/anaconda3/lib/python3.7/site-packages/pandas/core/reshape/merge.py:522: UserWarning: merging l
  warnings.warn(msg, UserWarning)
```

```
In [80]: m3 = mileage.copy()
         m3['date_str'] = m3.monthid.map(lambda x: str(x))
         m3['month'] = m3.date_str.str[-2:]
         m3['year'] = m3.date_str.str[:4]

In [81]: m3_2018 = (m3.loc[(m3.year == '2018') & (m3.State == 'california')]
         .groupby(['area']).agg({'Million_Vehicle_Miles':['sum']}))
         m3_2010 = (m3.loc[(m3.year == '2010') & (m3.State == 'california')]
         .groupby(['area']).agg({'Million_Vehicle_Miles':['sum']}))

In [82]: m3_2018.columns = ['miles']
         m3_2010.columns = ['miles']

In [83]: r2010 = m3_2010.iloc[1,][0]
         u2010 = m3_2010.iloc[2,][0]
         r2018 = m3_2018.iloc[1,][0]
         u2018 = m3_2018.iloc[2,][0]

In [84]: CA_2010_T = r2010 + u2010
         CA_2018_T = r2018 + u2018
         CA_Data = pd.DataFrame([[1,2,3,4],
                                 [CA_2010_T,(r2010+u2018),
                                  (r2010+u2018),CA_2018_T]]).T
         CA_Cover = pd.DataFrame([[2,3],
                                 [CA_2010_T,CA_2018_T]]).T

         CA_Data.columns = ['cat','miles']
         CA_Cover.columns = ['cat','miles']
         CA_Data.miles = CA_Data.miles.map(lambda x:x/1000)
         CA_Cover.miles = CA_Cover.miles.map(lambda x:x/1000)
```

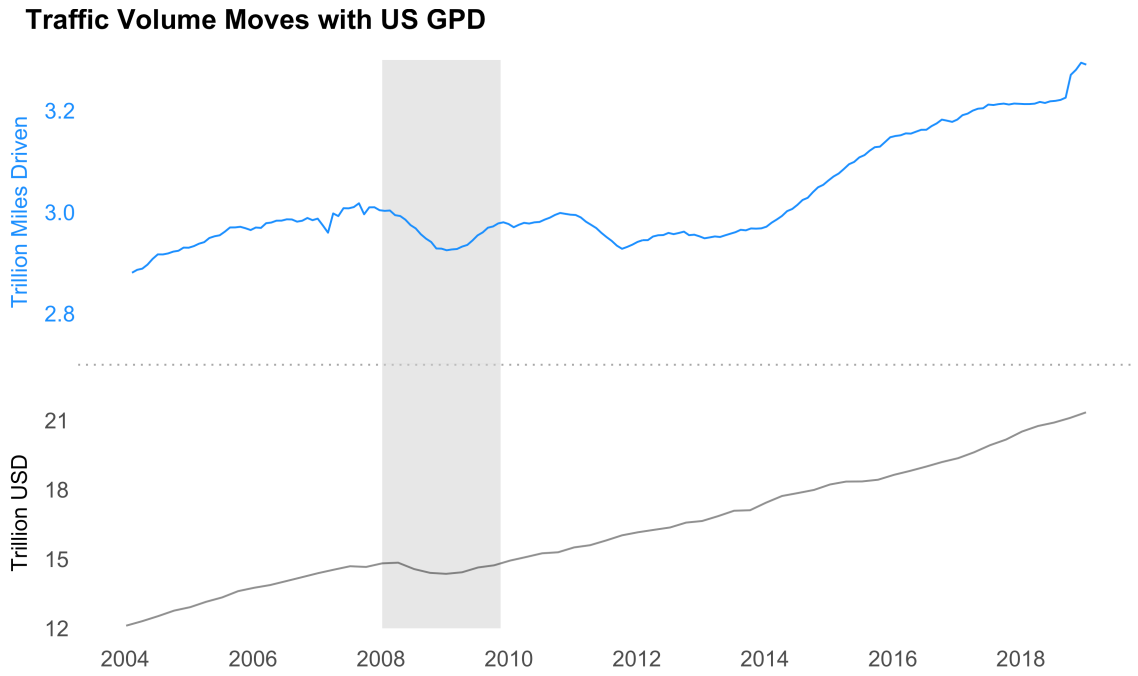
In [85]:
```
%R library(ggplot2)
%R library(gridExtra)
%R -i m1
%R -i gdp
%R -i mileage
%R -i ratios_2018
%R -i CA_Data
%R -i CA_Cover
```

In [86]:
```
%%R

p1 = (
ggplot()
+ geom_rect(aes(xmin=48.5,xmax=71.45,ymin=2.7,ymax=3.3),fill='lightgray',alpha=0.5)
+ geom_line(data=m1,aes(x=order,y=rolling),group=1,color='dodgerblue')
+ scale_x_continuous(breaks=seq(2,174,24),labels = seq(2004,2018,2))
+ ylab('Trillion Miles Driven')
+ theme_minimal()
+ theme(axis.title.x=element_blank())
+ theme(axis.text.x=element_blank())
+ theme(axis.title.y=element_text(size=12,color='dodgerblue',vjust=2.7))
+ theme(axis.text.y=element_text(color='dodgerblue',size=12))
+ theme(plot.margin = margin(t = 0, r = 0, b = -24.2, l = 5))
    # adjusting margin to line up graphs
+ scale_y_continuous(limits = c(2.7,3.3))
    # setting axis text so it won't overlap
+ theme(panel.grid=element_blank())
+ ggtitle("Traffic Volume Moves with US GPD")
+ theme(plot.title = element_text(hjust=-0.1,size=15,face='bold'))
)
```

In [87]:
```
%%R

p2 = (
ggplot()
+ theme_minimal()
+ scale_x_continuous(breaks=seq(3,60,8),labels = seq(2004,2018,2))
+ scale_y_continuous(breaks = seq(12,22,3),limits = c(12,23.5))
    # setting axis text so it won't overlap
+ theme(panel.grid = element_line('white'))
+ theme(axis.title.x = element_blank())
+ theme(axis.text.x = element_text(hjust=1.5,size=12))
+ theme(axis.title.y = element_text(size=12,hjust=0.4,vjust=2.7))
+ ylab("Trillion USD")
+ theme(axis.text.y = element_text(hjust=10,size=12))
+ theme(plot.margin = margin(t = 5, r = 0, b = 0, l = 5))
    # adjusting margin to line up graphs
+ theme(panel.grid=element_blank())
```

```
+ geom_hline(yintercept=23.4,color = 'darkgrey',linetype='dotted')
+ geom_rect(aes(xmin=16,xmax=23.4,ymin=12,ymax=23.44),fill='lightgray',alpha=0.5)
+ geom_line(data=gdp,aes(x=order,y=gpd_shrunk),group=1,alpha=0.5)
)
```

In [88]:
```
%%R -w 8.5 -h 5 --units in -r 400
grid.arrange(p1, p2, ncol = 1)
```

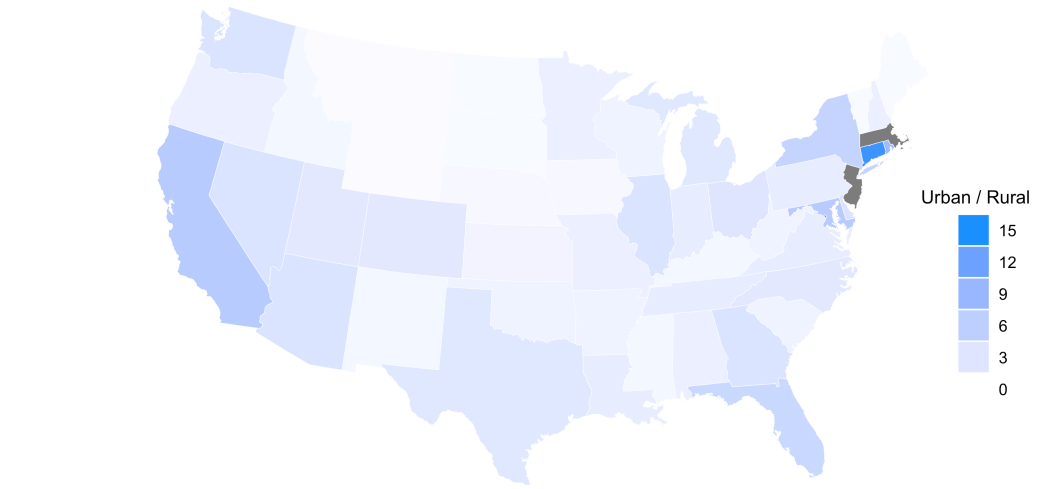


In [89]:
```
%%R
library(ggplot2)
library(dplyr)

us <- map_data("state")

map_2018 = (
ggplot()
+ geom_map(data=us, map=us,aes(x=long, y=lat, map_id=region),
           fill="#ffffff", color="#ffffff", size=0.15)
+ geom_map(data=ratios_2018, map=us,aes(fill=ratio, map_id=State),
           color="#ffffff", size=0.15)
+ scale_fill_continuous(name="Urban / Rural",
       limits = c(0,15),breaks=seq(0,15,3),
       low='white', high='dodgerblue',guide='colorbar')
+ theme(legend.title = element_text(size=4))
+ guides(fill = guide_legend(title = "Urban / Rural",reverse = TRUE))
+ theme(legend.title = element_text(margin = margin(l = -20), hjust = -15,size=10))
```

```
        + theme(legend.position = c(.9,0.40))
        + labs(x=NULL, y=NULL)
        + coord_map("albers", lat0 = 39, lat1 = 45)
        + theme(panel.border = element_blank())
        + theme(panel.background = element_blank())
        + theme(axis.ticks = element_blank())
        + theme(axis.text = element_blank())
        + ggtitle("In 2018 the Ratio of Urban to Rural Traffic Varies Greatly by State")
        + theme(plot.title = element_text(hjust=0,size=15,face='bold'))
        )
```

```
R[write to console]: Warning:
R[write to console]:  Ignoring unknown aesthetics: x, y
```

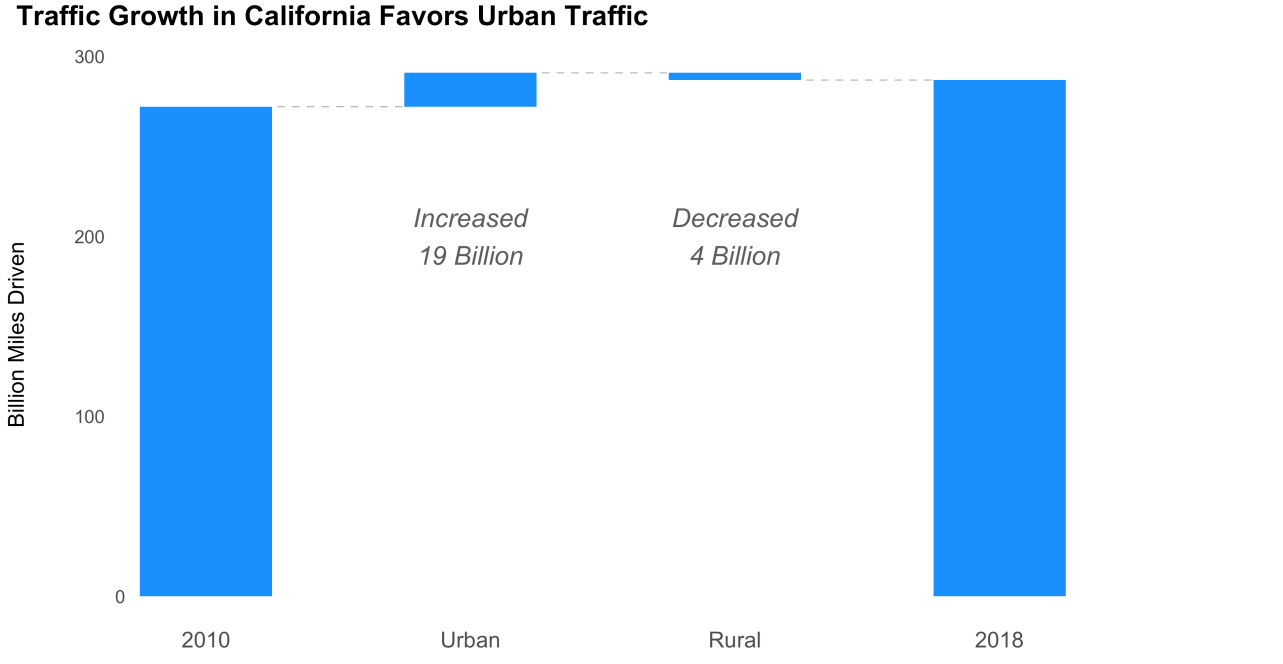
```
In [90]: %%R -w 8.5 -h 5 --units in --r 400
         map_2018
```




```
In [91]: %R ggsave('~/map_ratio.pdf',width=8.5,height=5)
```

```
In [92]: %%R -w 8.5 -h 5 --units in --r 400

         (
         ggplot()
         + geom_bar(data=CA_Data,aes(x=cat,y=miles),stat='identity',
```

```
          fill = 'dodgerblue',width=0.5)
+ geom_bar(data=CA_Cover,aes(x=cat,y=miles),stat='identity',fill ='white')
+ geom_segment(data=CA_Data,aes(x=1.27,xend=1.75,y=CA_Data[1,"miles"],
                                 yend=CA_Data[1,"miles"]),
                               linetype='dashed',color='darkgray',size=0.2)
+ geom_segment(data=CA_Data,aes(x=2.27,xend=2.75,y=CA_Data[2,"miles"],
                                 yend=CA_Data[2,"miles"]),
                               linetype='dashed',color='darkgray',size=0.2)
+ geom_segment(data=CA_Data,aes(x=3.27,xend=3.75,y=CA_Data[4,"miles"],
                                 yend=CA_Data[4,"miles"]),
                               linetype='dashed',color='darkgray',size=0.2)
+ geom_text(aes(x=2,y=200,label="Increased\n19 Billion"),
           size=5,alpha=0.7,fontface="italic")
+ geom_text(aes(x=3,y=200,label="Decreased\n4 Billion"),
           size=5,alpha=0.7,fontface='italic')
+ theme_minimal()
+ theme(panel.grid=element_blank())
+ scale_x_continuous(breaks=c(1,2,3,4),labels=c("2010","Urban","Rural","2018"))
+ theme(axis.text.x=element_text(size=12))
+ theme(axis.text.y = element_text(hjust=2,size=10))
+ scale_y_continuous(breaks=c(0,100,200,300),labels=c(0,100,200,300))
+ labs(y="Billion Miles Driven",x=NULL)
+ theme(axis.title.y = element_text(size=12))
+ ggtitle("Traffic Growth in California Favors Urban Traffic")
+ theme(plot.title = element_text(vjust=2,hjust=-.2,size=15,face='bold'))
)
```



```
In [93]: %R ggsave('~/waterfall.pdf',width=8.5,height=5)
```