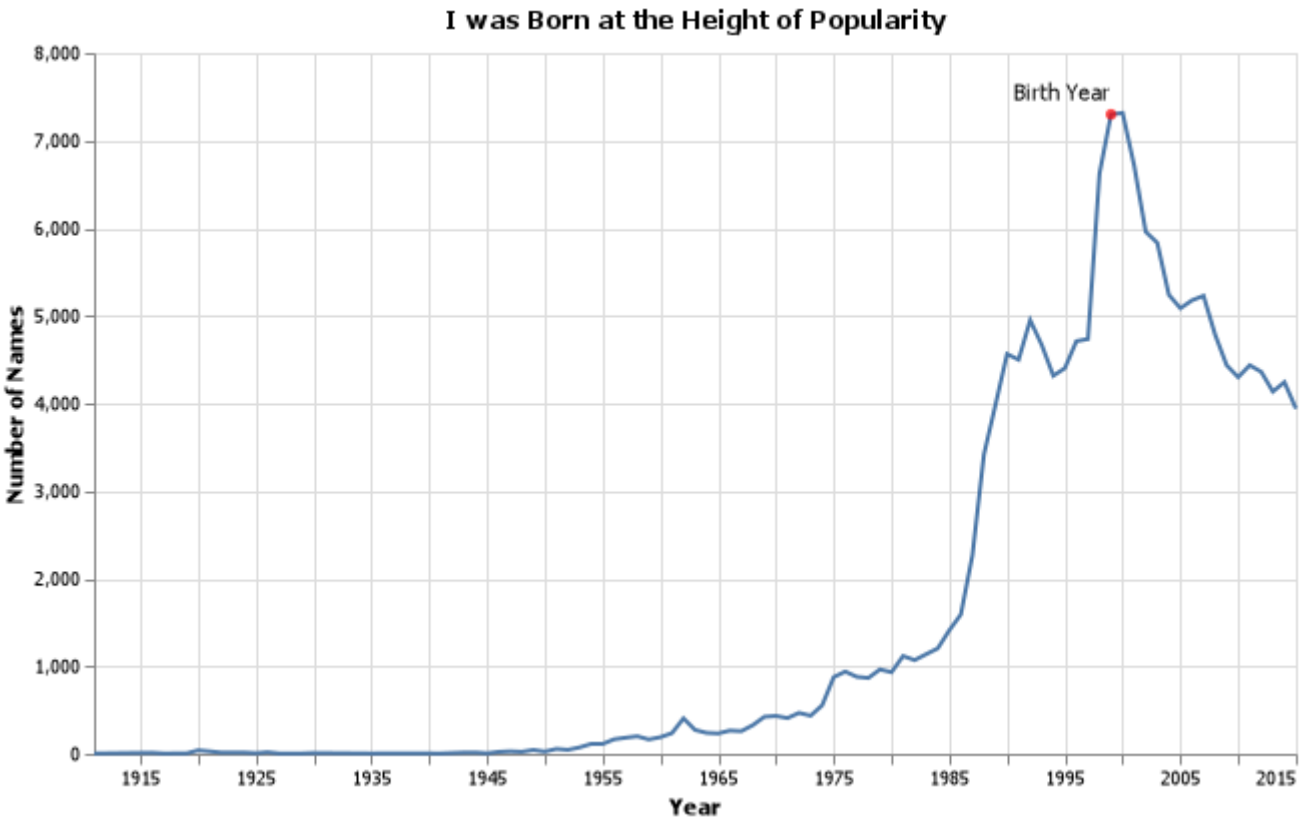
Week 2-3 What Is In a Name

Cameron Hansen

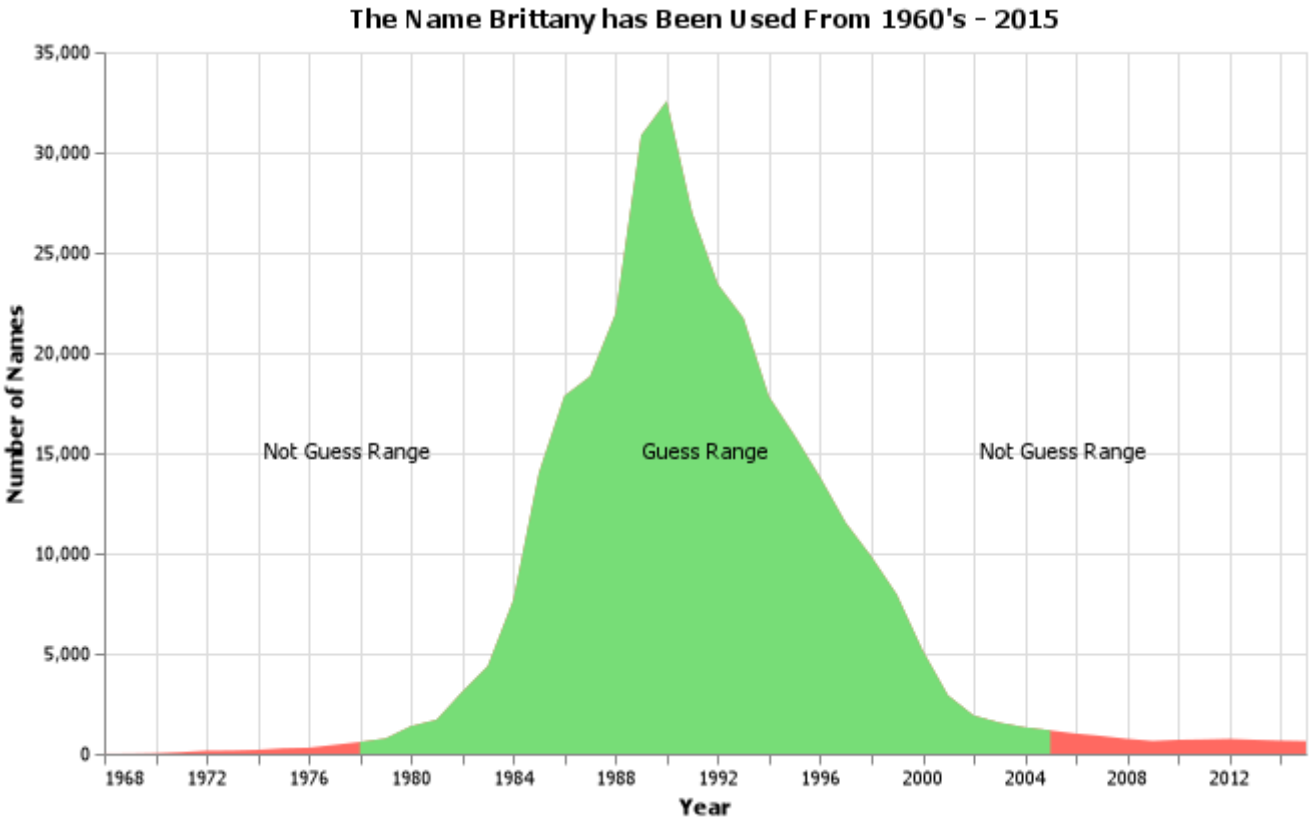
!(week2-3.py)

Elevator pitch

TECHNICAL DETAILS



The following chart is an analysis of my name,'Cameron', used from 1910 to 2015. The chart shows that I was born really at the peak of its usage.



Using the standard deviation of the year from a filtered dataset, I derived that most of the people named Brittany ranged from 16-43 years old.

- *YOU MIGHT INCLUDE SMALL CODE SNIPPETS TO HIGHLIGHT DECISIONS.*
- *YOU SHOULD HAVE QUALITY WRITING THAT DESCRIBES YOUR CHARTS AND TABLES.*

I HIGHLY RECOMMEND GRAMMARLY TO FIX YOUR SPELLING AND GRAMMAR. WRITING TAKES TIME TO BE CLEAR. SPEND THE TIME TO PRACITCE.

How does your name at your birth year compare to its use historically?

I was born at the height of popularity. My parents thought it was a good name that people couldn't make fun of, but it seems like people really liked it in general at the time. Makes sense why I have met so many Camerons in my life.

If you talked to someone named Brittany on the phone, what is your guess of their age? What ages would you not guess?

The name Brittany according to this dataset has only been used from 1960s to 2015 of which it has been popular from 1978 to 2005. If I were to guess what age it would be inbetween those ages. Given the year is 2021 their ages are inbetween 43 and 16 years old. I would definitaly not guess older than 51 years old.

Mary, Martha, Peter, and Paul are all Christian names. From 1920 - 2000, compare the name usage of each of the four names.

Think of a unique name from a famous movie. Plot that name and see how increases line up with the movie release.

...

APPENDIX A (PYTHON SCRIPT)

```
# %%
import pandas as pd
import altair as alt
import altair_saver
import numpy as np
alt.data_transformers.enable("json")
# %%
url = "https://raw.githubusercontent.com/byuidatascience/data4names/master/data-raw/names_year/r
names = pd.read_csv(url)
# %%
names.head()
names.year = pd.to_datetime(names.year, format='%Y')
names_agg = names.groupby('name')
names_agg.head()
my_name = names.query("name == 'Cameron'")
my_name.head()
# %%
cam_chart = (alt.Chart(my_name, title="I was Born at the Height of Popularity")
    .encode(
        alt.X('year(year):T', title = "Year"),
        alt.Y('Total:Q',title="Number of Names")
    )
    .mark_line().properties(width=600,height=350)
)

# %%
my_year = pd.DataFrame({
    'year' : [1999],
    'Total' : [my_name.query("year == 1999").Total.values[0]],
    'label' : ["Birth Year"]})

my_year.Total = my_year.Total.astype("int64",copy=True)
my_year.year = pd.to_datetime(my_year.year,format='%Y')

text_overlay = (alt.Chart(my_year).mark_text(align='right',dy=-10,baseline='middle')
    .encode
    (
        x = alt.X('year'),
        y = alt.Y('Total:Q'),
        text = 'label'
    )
)

my_point = (alt.Chart(my_year).mark_circle(color = 'red')
    .encode
    (
        x = alt.X('year'),
        y = alt.Y('Total:Q')
    )
)
```

```
cam_point = cam_chart  + text_overlay + my_point
# %%
cam_point
# %%
cam_point.save("Campoint.png")
# %%
# Brittany
brit = names.query("name == 'Brittany'")
brit_std = pd.DataFrame({
    'year' : [brit.year.median() + (brit.year.std() * 1),brit.year.median() - (brit.year.std() *
    'color' : ['red','green']
})



# %%
base = (alt.Chart(brit, title="The Name Brittany has Been Used From 1960's - 2015")
    .mark_area(color="#ff6961")
    .encode(
        x = alt.X('year', title="Year"),
        y = alt.Y('Total', title="Number of Names")
    )
    .properties(width=600,height=350)
)
# %%
area = (alt.Chart(brit.query("(year < @brit_std.year.values[0]) & (year > @brit_std.year.values[
    .mark_area(color='#77dd77')
    .encode(
        alt.X('year'),
        alt.Y('Total')
       )
    )



brit_final = base + area
# %%

brit_final.save("brit_final.png")
brit_final
# %%
# Christian names

christ_names = names.query("name == 'Mary' | name == 'Martha' |name == 'Peter' |name == 'Paul'")
christ_names = christ_names.query("((year >= 1920) & (year <= 2000))")
# %%
names_chart = (alt.Chart(christ_names, title="Christian are Closley Tied to Wars?")
    .encode(
        alt.X('year:T', title = "Year"),
        alt.Y('Total:Q',title="Number of Names"),
        color = 'name'
```

```
    )
    .mark_line()).properties(width=600,height=350)
# %%
wars = pd.DataFrame({
    'name' : ['World War II','Cold War', 'Korean War','Vietnam War', 'Gulf War'],
    'start' :[pd.to_datetime("1939"),pd.to_datetime("1947"),pd.to_datetime("1950"),pd.to_datetim
    'end' : [pd.to_datetime("1945"), pd.to_datetime("1991"),pd.to_datetime("1953"), pd.to_dateti
    'color' : ['black','#CD0000','#0047A0','#FFCD00',"#007A3D"],
    'y' : [33000,50000,45000,38000,10000]
})

name_overlay =(alt.Chart(wars)
    .mark_text(align='left',dy = -10)
    .encode(
    text = 'name',
    x = 'start',
    y = 'y'
    )
)
names_chart= names_chart + name_overlay

# %%
for war in wars.itertuples():
    new_data = christ_names.query("((year >= @war.start) & (year <= @war.end) & (name == 'Mary')
    war_chart = (alt.Chart(new_data)
    .mark_area(fill = war.color, opacity=0.4)
    .encode(
        x = 'year',
        y = 'Total',
    ))
    names_chart = names_chart + war_chart
# %%
names_chart.save("christ_names.png")
names_chart
# %%
```