```python
# To add a new cell, type '# %%'
# To add a new markdown cell, type '# %% [markdown]'
# %% [markdown]
# # Week 2 - Jacob Padgett
# %% [markdown]
# ## Grand Questions:
# *  How does your name at your birth year compare to its use historically?
# *  If you talked to someone named Brittany on the phone, what is your guess of their age? What ages would you not guess?
# *  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.

# %%
# Imports
import altair as alt
import calendar
import datetime
import numpy as np
import pandas as pd

# %% [markdown]
# ## Code for Question 1
# *  How does your name at your birth year compare to its use historically?

# %%
# Read in data
url = "https://github.com/byuidatascience/data4names/raw/master/data-raw/names_year/names_year.csv"
df = pd.read_csv(url)

all_jacob = df.query('name == "Jacob"')  # Narrow down to only Jacob's
jacob_1983 = df.query('name == "Jacob" & year == 1983')  # 538 Jacob's in 1983


# %%
all_jacob_CA_chart = (
    alt.Chart(all_jacob)
    .encode(x="year:O", y="CA")
    .mark_line()
    .properties(title="Q1. California Name Popularity - Jacob", width=800)
)
all_jacob_CA_chart.save("all_jacob_CA_chart.png")

# %% [markdown]
# ### Answered - Question 1
# *  How does your name at your birth year compare to its use historically?
# ---
#
# The name "Jacob" was used 538 (in CA where I was born) times in 1983 (my birth year) and here's a graph of how it has been used historically.
#
#
# ![](all_jacob_CA_chart.png)
# %% [markdown]
# ## Code for Question 2
# *  If you talked to someone named Brittany on the phone, what is your guess of their age? What ages would you not guess?

# %%
# See a visual of Brittany's birth years
guess_age_Brittany = df.query('name == "Brittany"')
guess_age_Brittany_chart = (
    alt.Chart(guess_age_Brittany)
    .encode(x="year:O", y="Total")
    .mark_line()
    .properties(title="Q2. Average Age of Brittany")
)
guess_age_Brittany_chart.save("guess_age_Brittany_chart.png")


# %%
# Helper Functions
def years_and_months(float_year):
    """Convert years with decimals into tuples of (year,month)"""
    year = int(float_year)
    month = int((float_year % 1) * 12)
    return year, month


def month_months(num):
    """To calculate if a month is plural or not"""
    if num == 1:
        return "month"
    else:
        return "months"


# %%
average_birth_year_for_Brittany = guess_age_Brittany.mean()[0]  # 1991.5
birth_year, birth_month = years_and_months(average_birth_year_for_Brittany)  # 1991,6

dt = datetime.datetime.today()  # Todays date for calculating age
current_year = dt.year  # 2021
current_month = dt.month  # 4

d0 = datetime.date(birth_year, birth_month, 1)  # 1991, 6, 1
d1 = datetime.date(current_year, current_month, 1)  # 2021, 4, 1
day_age = d1 - d0  # Day's old - 10897 and counting
average_age_year, average_age_month = years_and_months(
    day_age.days / 365
)  # (29, 10) and counting

print(
    f"""
The average Brittany was born in {calendar.month_name[birth_month]} of {birth_year}.
This would make the average Brittany {average_age_year} years and {average_age_month} {month_months(average_age_month)}, and that's how old I would guess she would be."""
)

# %% [markdown]
# ### Answered - Question 2
# *  If you talked to someone named Brittany on the phone, what is your guess of their age? What ages would you not guess?
# ---
```

```
#
# As mentioned in the output for the cell above:
# ```
# The average Brittany was born in June of 1991.
# This would make the average Brittany 29 years and 10 months, and that's how old I would guess she would be.
# ```
#
# Using the chart below, I would not guess Brittany is born prior to 1980, nor would Iguess she's born after 2001.
#
# ![](guess_age_Brittany_chart.png)
#
# %% [markdown]
# ## Code for Question 3
# *  Mary, Martha, Peter, and Paul are all Christian names. From 1920 - 2000, compare the name usage of each of the four names.

# %%
# Subset the data
mmpp = df.query('name in ["Mary","Martha","Peter","Paul"] & year > 1919 & year < 2020')


# %%
# Chart the subset
mmpp_chart = (
    alt.Chart(mmpp)
    .encode(alt.X("year:O"), alt.Y("Total:Q"), color="name")
    .mark_line()
    .properties(width=800, title="Q3. Mary, Martha, Peter & Paul by Year")
)
mmpp_chart.save("mmpp_chart.png")

# %% [markdown]
# ### Answered - Question 3
# * Mary, Martha, Peter, and Paul are all Christian names. From 1920 - 2000, compare the name usage of each of the four names.
# ---
#
# The following chart shows comparison between the names Mary, Martha, Peter & Paul between the years 1920 & 2020
#
# ![](mmpp_chart.png)
# %% [markdown]
# ## Code for Question 4
# *  Think of a unique name from a famous movie. Plot that name and see how increases line up with the movie release.

# %%
titanic_Jack = df.query(
    'name == "Jack" & year >= 1987 & year <= 2007'
)  # From the Titanic in 1997
titanic_Jack_chart = (
    alt.Chart(titanic_Jack)
    .encode(x="year:O", y="Total")
    .mark_line()
    .properties(title="Q4. Jack From Titanic")
)  # Build chart
titanic_Jack_chart.save("titanic_Jack_chart.png")  # Save chart
```

```

# %% [markdown]
# ### Answered - Question 4
# * Think of a unique name from a famous movie. Plot that name and see how increases line up with the movie release.
# ---
#
# With the movie Titanic being released in the year 1997, one of the two main characters, Jack is who I chose to evaluate. The name Jack was already on the up-trend when the movie came out, and it didn't hurt it. In fact, the name kept gaining popularity for about 8 years after..
#
# See the below chart for details.
#
# ![](titanic_Jack_chart.png)

```