The Nobel Prize has been among the most prestigious international awards since 1901. Each year, awards are bestowed in chemistry, literature, physics, physiology or medicine, economics, and peace. In addition to the honor, prestige, and substantial prize money, the recipient also gets a gold medal with an image of Alfred Nobel (1833 - 1896), who established the prize.

The Nobel Foundation has made a dataset available of all prize winners from the outset of the awards from 1901 to 2023. The dataset used in this project is from the Nobel Prize API and is available in the `nobel.csv` file in the `data` folder.

In this project, you'll get a chance to explore and answer several questions related to this prizewinning data. And we encourage you then to explore further questions that you're interested in!

```
# Loading in required libraries
import pandas as pd
import seaborn as sns
import numpy as np
import matplotlib.pyplot as plt

# Start coding here!
nobel_df = pd.read_csv(r'data/nobel.csv')
nobel_df.tail(10)
```

		c…	prize	motivation	pri…	lau…
990	2023	Medicine	The Nobel Prize in Physiology or Medicine 20…	"for their discoveries concerning nucleoside …	1/2	102
991	2023	Physics	The Nobel Prize in Physics 2023	"for experimental methods that generate att…	1/3	102
992	2023	Physics	The Nobel Prize in Physics 2023	"for experimental methods that generate att…	1/3	102
993	2023	Physics	The Nobel Prize in Physics 2023	"for experimental methods that generate att…	1/3	102
994	2023	Chemistry	The Nobel Prize in Chemistry 2023	"for the discovery and synthesis of quantum …	1/3	102
995	2023	Chemistry	The Nobel Prize in Chemistry 2023	"for the discovery and synthesis of quantum …	1/3	103
996	2023	Chemistry	The Nobel Prize in Chemistry 2023	"for the discovery and synthesis of quantum …	1/3	103
997	2023	Literature	The Nobel Prize in Literature 2023	"for his innovative plays and prose which giv…	1/1	103
998	2023	Peace	The Nobel Peace Prize 2023	"for her fight against the oppression of wom…	1/1	103
999	2023	Economics	The Sveriges Riksbank Prize in Economic Sci…	"for having advanced our understanding of …	1/1	103

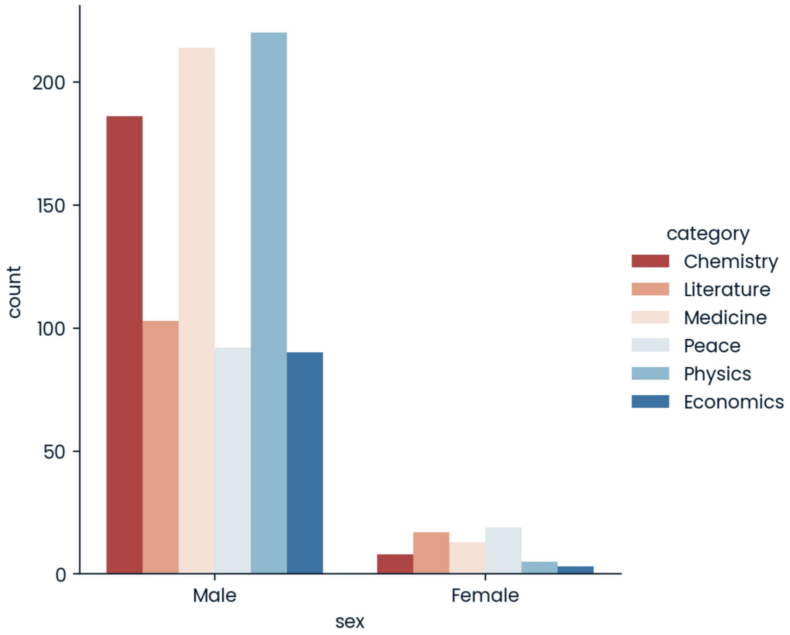
Rows: 10

```
# Most commonly awarded gender and birth country

top_gender = nobel_df['sex'].value_counts().idxmax()
without_sex = nobel_df[nobel_df['sex'].isnull()] # Actually they are Organization!

top_country = nobel_df['birth_country'].value_counts().idxmax()

# Visualisation
sns.set_palette('RdBu')
sns.catplot(data=nobel_df,
            x="sex",
            kind='count',
           hue= 'category')
plt.show()
```

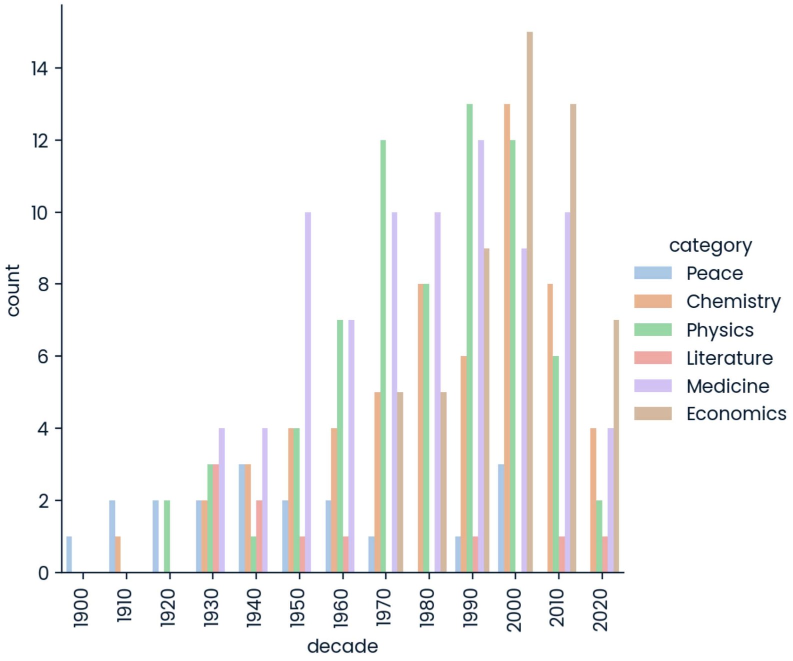


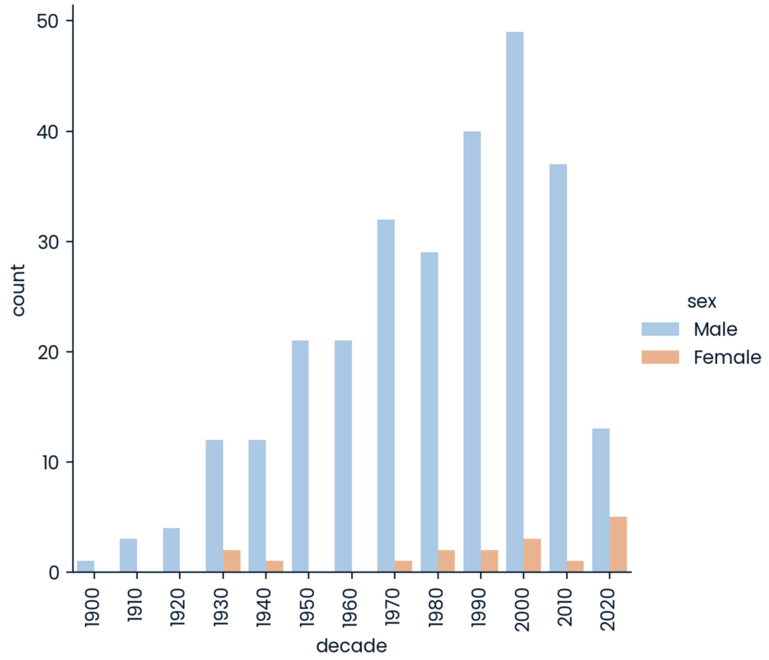
```
# Best decade awarded for USA

us_born = nobel_df.query("birth_country == 'United States of America'")
us_born['decade'] = us_born['year'] // 10 * 10
max_decade_usa = us_born['decade'].value_counts().idxmax()

# Visualisation
sns.set_palette("pastel")
sns.catplot(x="decade",
           data=us_born,
            kind='count',
           hue='category')
plt.xticks(rotation=90)

sns.catplot(x="decade",
           data=us_born,
            kind='count',
           hue='sex')
plt.xticks(rotation=90)
plt.show()
```



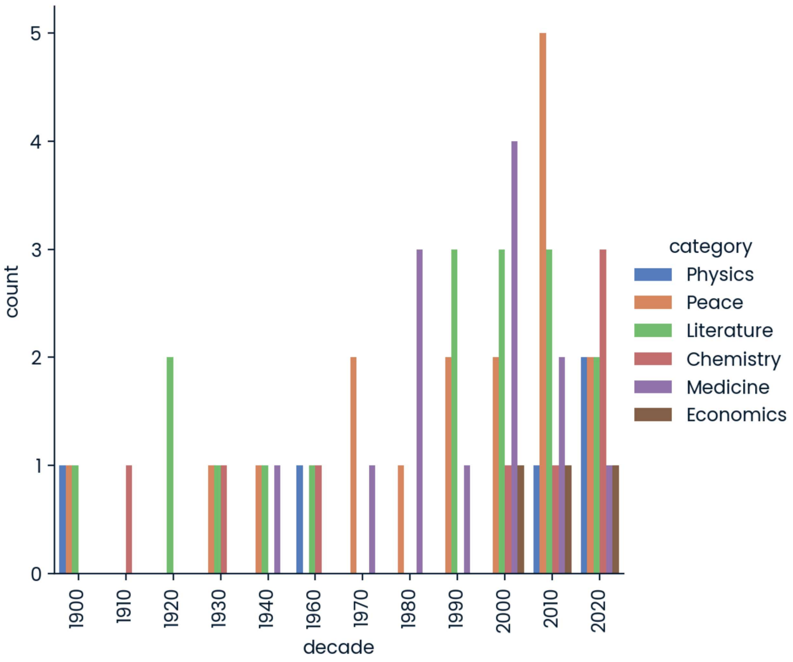


```
nobel_df['decade'] = nobel_df['year'] // 10 * 10  # Adding a decade column for easier accessing

# Filter for female laureates
female_laureates = nobel_df[nobel_df['sex'] == 'Female'][nobel_df['laureate_type'] =='Individual']

# Plotting the best decade for female laureates by category
sns.set_palette("muted")
sns.catplot(data=female_laureates,
            x="decade",
            kind='count',
            hue='category')
plt.xticks(rotation=90)
plt.show()


# Based on the plot we have:
max_female_dict = {2020: 'peace'}
```



```
first_women = nobel_df[nobel_df["sex"] == "Female"].iloc[0, :]
first_women_name = first_women["full_name"]
first_women_category = first_women["category"]
```

```
condition = nobel_df["full_name"].value_counts()
repeat_list = condition[condition>=2].index.tolist()
```