Let's see what patterns we can find in the data of the past Nobel laureates. What can we learn about the Nobel prize and our world more generally?

In [1]:

```
import pandas as pd
import numpy as np
import plotly.express as px
import seaborn as sns
import matplotlib.pyplot as plt
```

In [2]:

```
pd.options.display.float_format = '{:,.2f}'.format
```

In [3]:

```
df_data = pd.read_csv('nobel_prize_data.csv')
```

Data Exploration & Cleaning

In [4]:

```
df_data.shape
```

Out[4]:

```
(962, 16)
```

In [5]:

```
df_data.head()
```

Out[5]:

	year	category	prize	motivation	prize_share	laureate_type	full_name	birth_date
0	1901	Chemistry	The Nobel Prize in Chemistry 1901	"in recognition of the extraordinary services ...	1/1	Individual	Jacobus Henricus van 't Hoff	1852-08-30
1	1901	Literature	The Nobel Prize in Literature 1901	"in special recognition of his poetic composit...	1/1	Individual	Sully Prudhomme	1839-03-16
2	1901	Medicine	The Nobel Prize in Physiology or Medicine 1901	"for his work on serum therapy, especially its...	1/1	Individual	Emil Adolf von Behring	1854-03-15
3	1901	Peace	The Nobel Peace Prize 1901	NaN	1/2	Individual	Frédéric Passy	1822-05-20
4	1901	Peace	The Nobel Peace Prize 1901	NaN	1/2	Individual	Jean Henry Dunant	1828-05-08

In [6]:

```
df_data.tail()
```

Out[6]:

	year	category	prize	motivation	prize_share	laureate_type	full_name	birth_dat
957	2020	Medicine	The Nobel Prize in Physiology or Medicine 2020	“for the discovery of Hepatitis C virus”	1/3	Individual	Michael Houghton	1949-07 02
958	2020	Peace	The Nobel Peace Prize 2020	“for its efforts to combat hunger, for its con...	1/1	Organization	World Food Programme (WFP)	NaN
959	2020	Physics	The Nobel Prize in Physics 2020	“for the discovery of a supermassive compact o...	1/4	Individual	Andrea Ghez	1965-06 16
960	2020	Physics	The Nobel Prize in Physics 2020	“for the discovery of a supermassive compact o...	1/4	Individual	Reinhard Genzel	1952-03 24
961	2020	Physics	The Nobel Prize in Physics 2020	“for the discovery that black hole formation i...	1/2	Individual	Roger Penrose	1931-08 08

In [7]:

```
print(df_data.duplicated().values.any())
```

```
False
```

In [8]:

```
print(df_data.isna().values.any())
```

```
True
```

In [9]:

```
df_data.isna().sum()
```

Out[9]:

```
year                       0
category                   0
prize                      0
motivation                88
prize_share                0
laureate_type              0
full_name                  0
birth_date                28
birth_city                31
birth_country             28
birth_country_current     28
sex                       28
organization_name        255
organization_city        255
organization_country     254
ISO                       28
dtype: int64
```

In [10]:

```
col_subset = ['year','category', 'laureate_type',
              'birth_date','full_name', 'organization_name']
df_data.loc[df_data.birth_date.isna()][col_subset]
```

Out[10]:

	year	category	laureate_type	birth_date	full_name	organization_name
24	1904	Peace	Organization	NaN	Institut de droit international (Institute of ...	NaN
60	1910	Peace	Organization	NaN	Bureau international permanent de la Paix (Per...	NaN
89	1917	Peace	Organization	NaN	Comité international de la Croix Rouge (Intern...	NaN
200	1938	Peace	Organization	NaN	Office international Nansen pour les Réfugiés ...	NaN
215	1944	Peace	Organization	NaN	Comité international de la Croix Rouge (Intern...	NaN
237	1947	Peace	Organization	NaN	American Friends Service Committee (The Quakers)	NaN
238	1947	Peace	Organization	NaN	Friends Service Council (The Quakers)	NaN
283	1954	Peace	Organization	NaN	Office of the United Nations High Commissioner...	NaN
348	1963	Peace	Organization	NaN	Comité international de la Croix Rouge (Intern...	NaN
349	1963	Peace	Organization	NaN	Ligue des Sociétés de la Croix-Rouge (League o...	NaN
366	1965	Peace	Organization	NaN	United Nations Children's Fund (UNICEF)	NaN
399	1969	Peace	Organization	NaN	International Labour Organization (I.L.O.)	NaN
479	1977	Peace	Organization	NaN	Amnesty International	NaN
523	1981	Peace	Organization	NaN	Office of the United Nations High Commissioner...	NaN
558	1985	Peace	Organization	NaN	International Physicians for the Prevention of...	NaN
588	1988	Peace	Organization	NaN	United Nations Peacekeeping Forces	NaN
659	1995	Peace	Organization	NaN	Pugwash Conferences on Science and World Affairs	NaN
682	1997	Peace	Organization	NaN	International Campaign to Ban Landmines (ICBL)	NaN
703	1999	Peace	Organization	NaN	Médecins Sans Frontières	NaN
730	2001	Peace	Organization	NaN	United Nations (U.N.)	NaN
778	2005	Peace	Organization	NaN	International Atomic Energy Agency (IAEA)	NaN
788	2006	Peace	Organization	NaN	Grameen Bank	NaN
801	2007	Peace	Organization	NaN	Intergovernmental Panel on Climate Change (IPCC)	NaN

	year	category	laureate_type	birth_date	full_name	organization_name
860	2012	Peace	Organization	NaN	European Union (EU)	NaN
873	2013	Peace	Organization	NaN	Organisation for the Prohibition of Chemical W...	NaN
897	2015	Peace	Organization	NaN	National Dialogue Quartet	NaN
919	2017	Peace	Organization	NaN	International Campaign to Abolish Nuclear Weap...	NaN
958	2020	Peace	Organization	NaN	World Food Programme (WFP)	NaN

In [11]:

```
col_subset = ['year','category', 'laureate_type','full_name', 'organization_name']
df_data.loc[df_data.organization_name.isna()][col_subset]
```

Out[11]:

	year	category	laureate_type	full_name	organization_name
1	1901	Literature	Individual	Sully Prudhomme	NaN
3	1901	Peace	Individual	Frédéric Passy	NaN
4	1901	Peace	Individual	Jean Henry Dunant	NaN
7	1902	Literature	Individual	Christian Matthias Theodor Mommsen	NaN
9	1902	Peace	Individual	Charles Albert Gobat	NaN
...	...	...	...	...	...
932	2018	Peace	Individual	Nadia Murad	NaN
942	2019	Literature	Individual	Peter Handke	NaN
946	2019	Peace	Individual	Abiy Ahmed Ali	NaN
954	2020	Literature	Individual	Louise Glück	NaN
958	2020	Peace	Organization	World Food Programme (WFP)	NaN

255 rows × 5 columns

In [12]:

```
df_data.birth_date = pd.to_datetime(df_data.birth_date)
```

In [13]:

```
separated_values = df_data.prize_share.str.split('/', expand=True)
numerator = pd.to_numeric(separated_values[0])
denomenator = pd.to_numeric(separated_values[1])
df_data['share_pct'] = numerator / denomenator
```

In [14]:

```
df_data.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 962 entries, 0 to 961
Data columns (total 17 columns):
 #   Column                 Non-Null Count  Dtype         
---  ------                 --------------  -----         
 0   year                   962 non-null    int64         
 1   category               962 non-null    object        
 2   prize                  962 non-null    object        
 3   motivation             874 non-null    object        
 4   prize_share            962 non-null    object        
 5   laureate_type          962 non-null    object        
 6   full_name              962 non-null    object        
 7   birth_date             934 non-null    datetime64[ns]
 8   birth_city             931 non-null    object        
 9   birth_country          934 non-null    object        
 10  birth_country_current  934 non-null    object        
 11  sex                    934 non-null    object        
 12  organization_name      707 non-null    object        
 13  organization_city      707 non-null    object        
 14  organization_country   708 non-null    object        
 15  ISO                    934 non-null    object        
 16  share_pct              962 non-null    float64       
dtypes: datetime64[ns](1), float64(1), int64(1), object(14)
memory usage: 127.9+ KB
```

Who were the first 3 Women to Win the Nobel Prize?

- What are the names of the first 3 female Nobel laureates?
- What did the win the prize for?

In [15]:

```
df_data[df_data.sex == 'Female'].sort_values('year', ascending=True)[:3]
```

Out[15]:

	year	category	prize	motivation	prize_share	laureate_type	full_name	birth_date
18	1903	Physics	The Nobel Prize in Physics 1903	"in recognition of the extraordinary services ...	1/4	Individual	Marie Curie, née Sklodowska	1867-11-07
29	1905	Peace	The Nobel Peace Prize 1905	NaN	1/1	Individual	Baroness Bertha Sophie Felicita von Suttner, n...	1843-06-09
51	1909	Literature	The Nobel Prize in Literature 1909	"in appreciation of the lofty idealism, vivid ...	1/1	Individual	Selma Ottilia Lovisa Lagerlöf	1858-11-20

Find the Repeat Winners

In [16]:

```
is_winner = df_data.duplicated(subset=['full_name'], keep=False)
multiple_winners = df_data[is_winner]
print(f'There are {multiple_winners.full_name.nunique()}' \
      ' winners who were awarded the prize more than once.')
```

```
There are 6 winners who were awarded the prize more than once.
```

In [17]:

```
col_subset = ['year', 'category', 'laureate_type', 'full_name']
multiple_winners[col_subset]
```

Out[17]:

	year	category	laureate_type	full_name
18	1903	Physics	Individual	Marie Curie, née Sklodowska
62	1911	Chemistry	Individual	Marie Curie, née Sklodowska
89	1917	Peace	Organization	Comité international de la Croix Rouge (Intern...
215	1944	Peace	Organization	Comité international de la Croix Rouge (Intern...
278	1954	Chemistry	Individual	Linus Carl Pauling
283	1954	Peace	Organization	Office of the United Nations High Commissioner...
297	1956	Physics	Individual	John Bardeen
306	1958	Chemistry	Individual	Frederick Sanger
340	1962	Peace	Individual	Linus Carl Pauling
348	1963	Peace	Organization	Comité international de la Croix Rouge (Intern...
424	1972	Physics	Individual	John Bardeen
505	1980	Chemistry	Individual	Frederick Sanger
523	1981	Peace	Organization	Office of the United Nations High Commissioner...

Number of Prizes per Category

- In how many categories are prizes awarded?
- Which category has the most number of prizes awarded?
- Which category has the fewest number of prizes awarded?

In [18]:

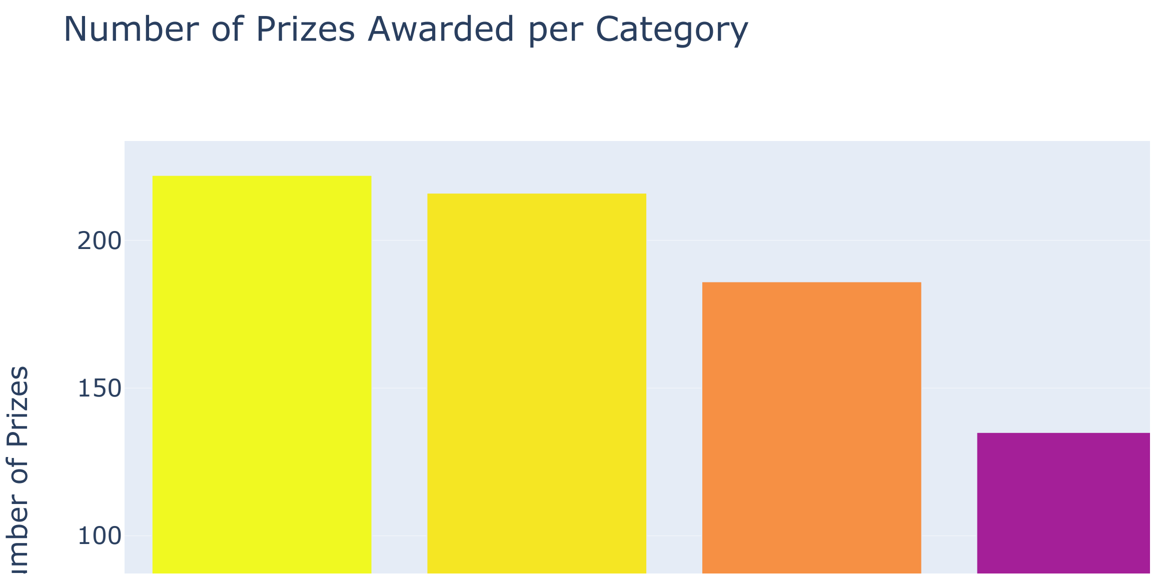
```
df_data.category.nunique()
```

Out[18]:

```
6
```

In [19]:

```
prizes_per_category = df_data.category.value_counts()
v_bar = px.bar(
        x = prizes_per_category.index,
        y = prizes_per_category.values,
        color = prizes_per_category.values,
        title='Number of Prizes Awarded per Category')

v_bar.update_layout(xaxis_title='Nobel Prize Category',
                    coloraxis_showscale=False,
                    yaxis_title='Number of Prizes')
v_bar.show()
```



- When was the first prize in the field of Economics awarded?
- Who did the prize go to?

In [20]:

```
df_data[df_data.category == 'Economics'].sort_values('year')[:3]
```

Out[20]:

	year	category	prize	motivation	prize_share	laureate_type	full_name	birth_date
393	1969	Economics	The Sveriges Riksbank Prize in Economic Scienc...	"for having developed and applied dynamic mode...	1/2	Individual	Jan Tinbergen	1903-04-12
394	1969	Economics	The Sveriges Riksbank Prize in Economic Scienc...	"for having developed and applied dynamic mode...	1/2	Individual	Ragnar Frisch	1895-03-03
402	1970	Economics	The Sveriges Riksbank Prize in Economic Scienc...	"for the scientific work through which he has ...	1/1	Individual	Paul A. Samuelson	1915-05-15

Male and Female Winners by Category

In [21]:

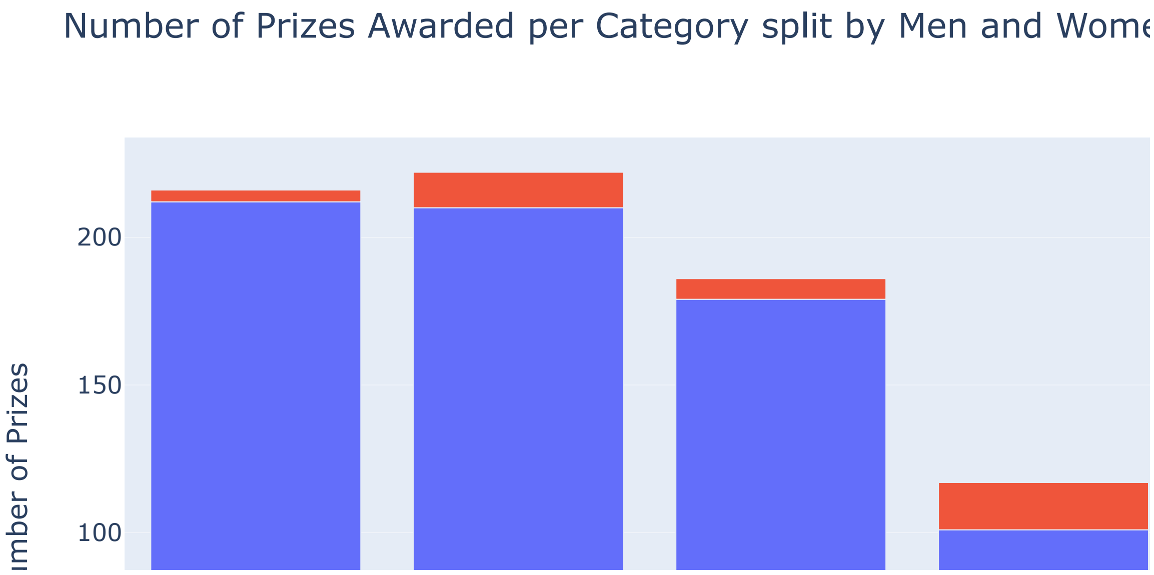
```
cat_men_women = df_data.groupby(['category', 'sex'],
                               as_index=False).agg({'prize': pd.Series.count})
cat_men_women.sort_values('prize', ascending=False, inplace=True)
cat_men_women
```

Out[21]:

```
      category     sex  prize
11     Physics    Male    212
 7    Medicine    Male    210
 1   Chemistry    Male    179
 5  Literature    Male    101
 9       Peace    Male     90
 3   Economics    Male     84
 8       Peace  Female     17
 4  Literature  Female     16
 6    Medicine  Female     12
 0   Chemistry  Female      7
10     Physics  Female      4
 2   Economics  Female      2
```

In [22]:

```
v_bar_split = px.bar(x = cat_men_women.category,
                     y = cat_men_women.prize,
                     color = cat_men_women.sex,
                     title='Number of Prizes Awarded per Category split by Men and Women')

v_bar_split.update_layout(xaxis_title='Nobel Prize Category',
                          yaxis_title='Number of Prizes')
v_bar_split.show()
```



Number of Prizes Awarded Over Time

Are more prizes awarded recently than when the prize was first created? Show the trend in awards visually.

In [23]:

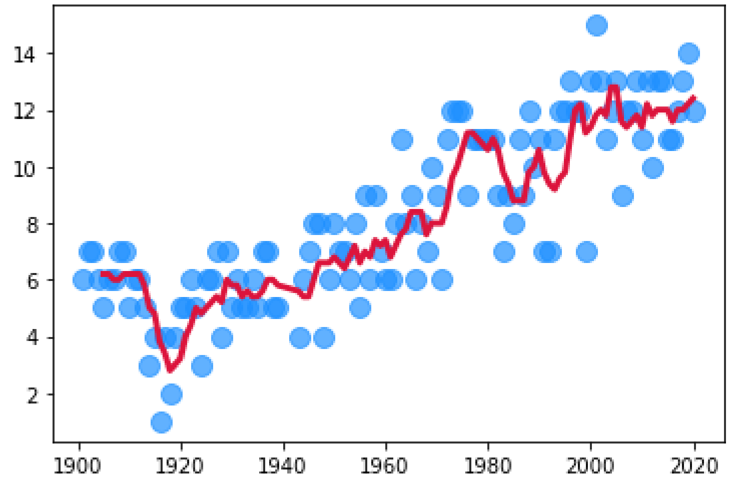
```
prize_per_year = df_data.groupby(by='year').count().prize
```

In [24]:

```
moving_average = prize_per_year.rolling(window=5).mean()
```

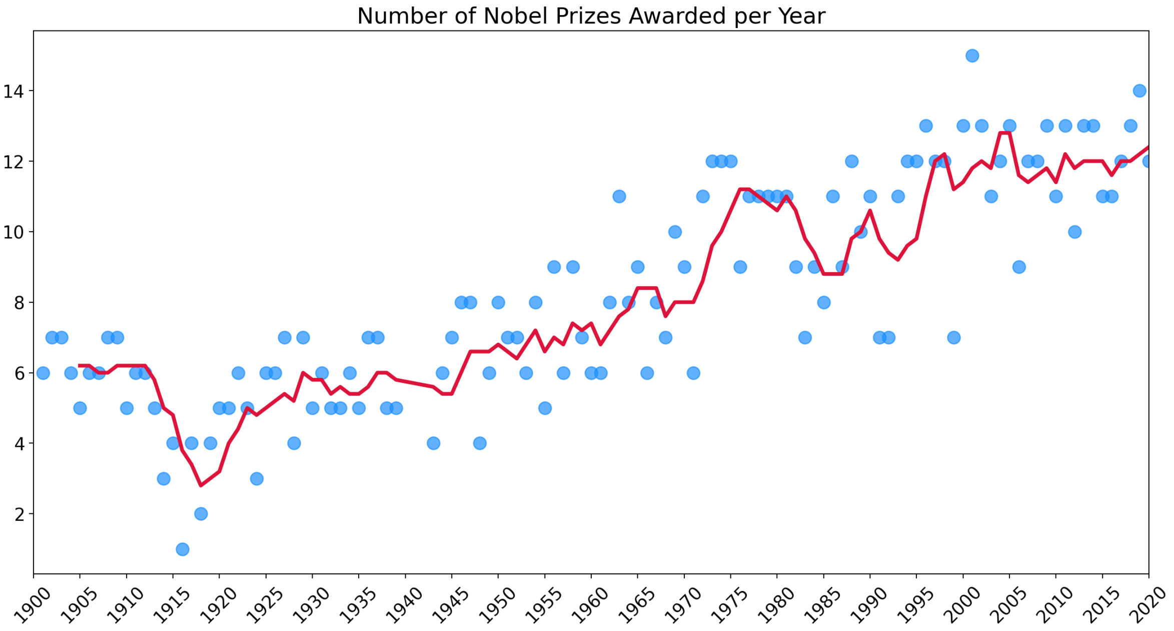
In [25]:

```
plt.scatter(x=prize_per_year.index,
            y=prize_per_year.values,
            c='dodgerblue',
            alpha=0.7,
            s=100,)

plt.plot(prize_per_year.index,
         moving_average.values,
         c='crimson',
         linewidth=3,)

plt.show()
```



In [26]:

```
plt.figure(figsize=(16,8), dpi=200)
plt.title('Number of Nobel Prizes Awarded per Year', fontsize=18)
plt.yticks(fontsize=14)
plt.xticks(ticks=np.arange(1900, 2021, step=5),
           fontsize=14,
           rotation=45)

ax = plt.gca()
ax.set_xlim(1900, 2020)

ax.scatter(x=prize_per_year.index,
           y=prize_per_year.values,
           c='dodgerblue',
           alpha=0.7,
           s=100,)

ax.plot(prize_per_year.index,
        moving_average.values,
        c='crimson',
        linewidth=3,)

plt.show()
```



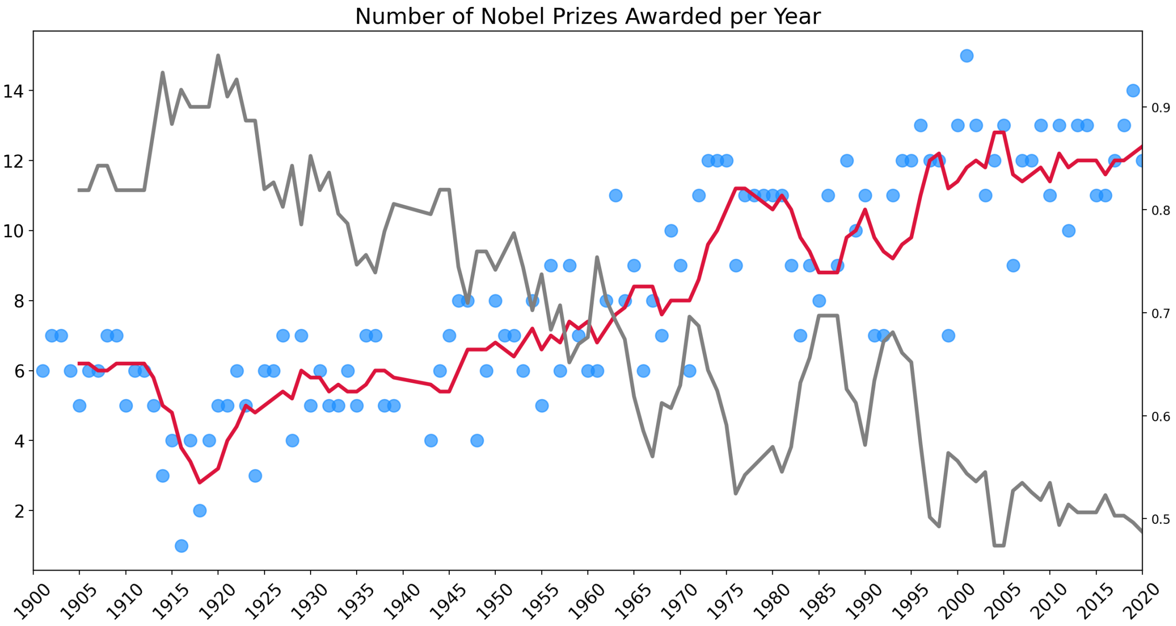
Are More Prizes Shared Than Before?

In [27]:

```
yearly_avg_share = df_data.groupby(by='year').agg({'share_pct': pd.Series.mean})
share_moving_average = yearly_avg_share.rolling(window=5).mean()
```

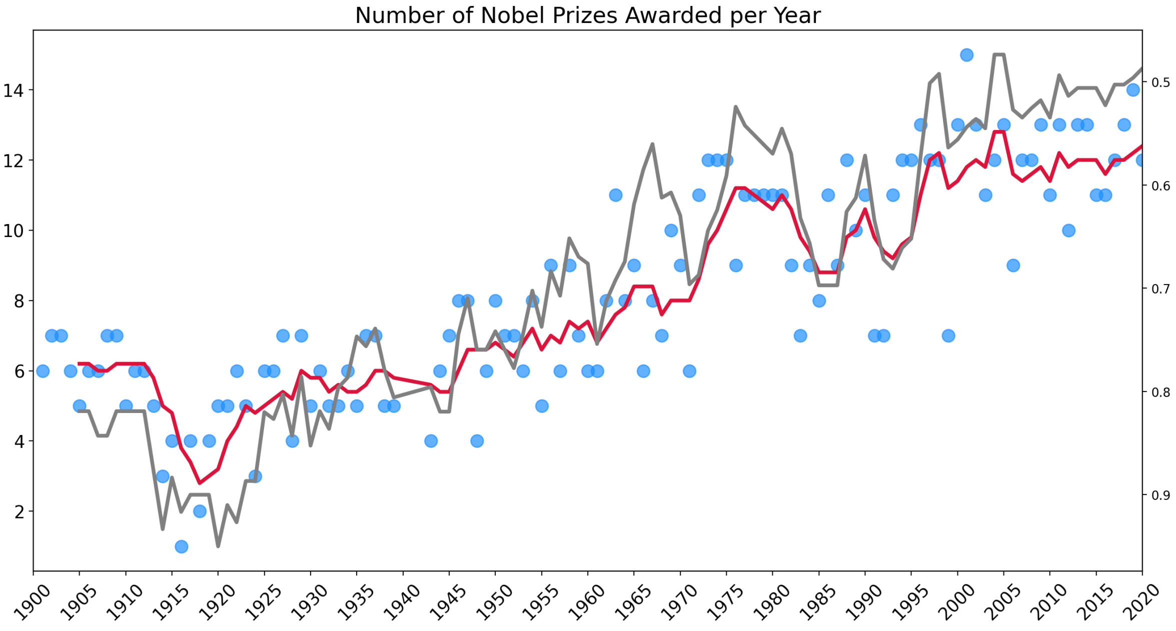
In [28]:

```
plt.figure(figsize=(16,8), dpi=200)
plt.title('Number of Nobel Prizes Awarded per Year', fontsize=18)
plt.yticks(fontsize=14)
plt.xticks(ticks=np.arange(1900, 2021, step=5),
           fontsize=14,
           rotation=45)

ax1 = plt.gca()
ax2 = ax1.twinx()
ax1.set_xlim(1900, 2020)

ax1.scatter(x=prize_per_year.index,
           y=prize_per_year.values,
           c='dodgerblue',
           alpha=0.7,
           s=100,)

ax1.plot(prize_per_year.index,
        moving_average.values,
        c='crimson',
        linewidth=3,)

ax2.plot(prize_per_year.index,
        share_moving_average.values,
        c='grey',
        linewidth=3,)

plt.show()
```



In [29]:

```
plt.figure(figsize=(16,8), dpi=200)
plt.title('Number of Nobel Prizes Awarded per Year', fontsize=18)
plt.yticks(fontsize=14)
plt.xticks(ticks=np.arange(1900, 2021, step=5),
           fontsize=14,
           rotation=45)

ax1 = plt.gca()
ax2 = ax1.twinx()
ax1.set_xlim(1900, 2020)


ax2.invert_yaxis()

ax1.scatter(x=prize_per_year.index,
           y=prize_per_year.values,
           c='dodgerblue',
           alpha=0.7,
           s=100,)

ax1.plot(prize_per_year.index,
        moving_average.values,
        c='crimson',
        linewidth=3,)

ax2.plot(prize_per_year.index,
        share_moving_average.values,
        c='grey',
        linewidth=3,)

plt.show()
```



The Countries with the Most Nobel Prizes

- What is the ranking for the top 20 countries in terms of the number of prizes?

In [30]:

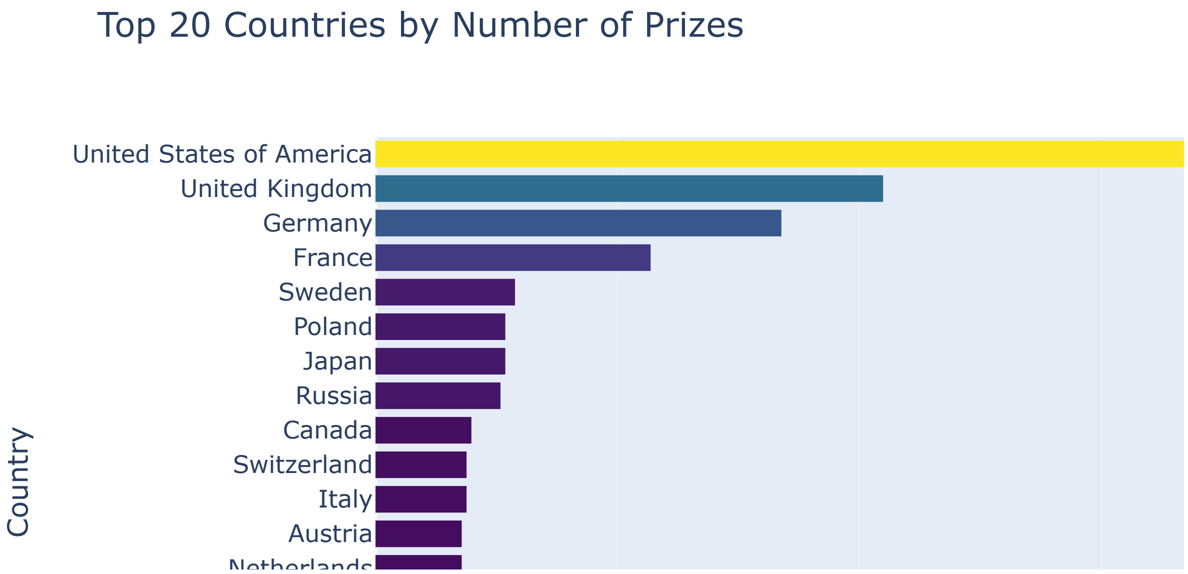
```
top_countries = df_data.groupby(['birth_country_current'],
                                  as_index=False).agg({'prize': pd.Series.count})

top_countries.sort_values(by='prize', inplace=True)
top20_countries = top_countries[-20:]
top20_countries
```

Out[30]:

	birth_country_current	prize
7	Belgium	9
31	Hungary	9
33	India	9
2	Australia	10
20	Denmark	12
54	Norway	12
13	China	12
51	Netherlands	18
3	Austria	18
39	Italy	19
68	Switzerland	19
11	Canada	20
61	Russia	26
40	Japan	27
57	Poland	27
67	Sweden	29
25	France	57
26	Germany	84
73	United Kingdom	105
74	United States of America	281

In [31]:

```
h_bar = px.bar(x=top20_countries.prize,
               y=top20_countries.birth_country_current,
               orientation='h',
               color=top20_countries.prize,
               color_continuous_scale='Viridis',
               title='Top 20 Countries by Number of Prizes')

h_bar.update_layout(xaxis_title='Number of Prizes',
                    yaxis_title='Country',
                    coloraxis_showscale=False)
h_bar.show()
```



Choropleth Map to Show the Number of Prizes Won by Country

In [32]:

```
df_countries = df_data.groupby(['birth_country_current', 'ISO'],
                               as_index=False).agg({'prize': pd.Series.count})
df_countries.sort_values('prize', ascending=False)
```

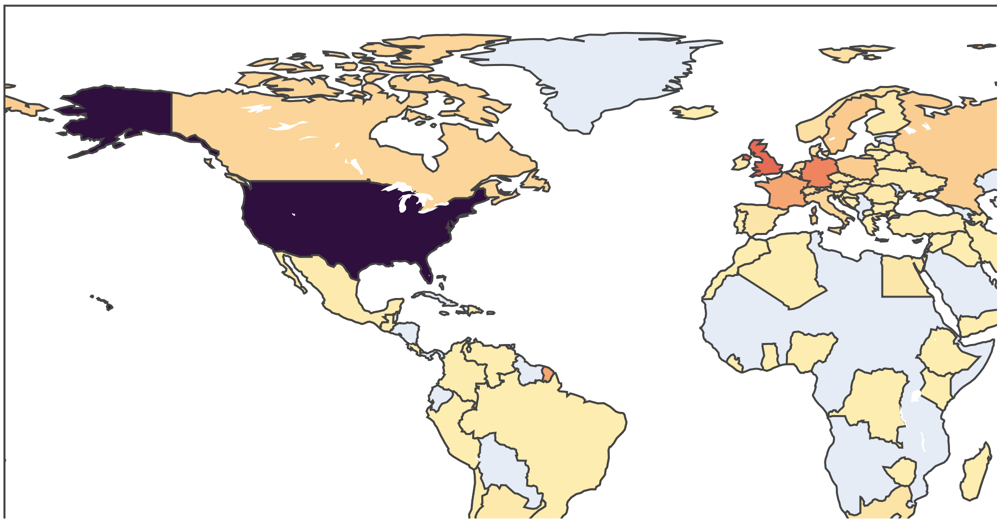
Out[32]:

	birth_country_current	ISO	prize
74	United States of America	USA	281
73	United Kingdom	GBR	105
26	Germany	DEU	84
25	France	FRA	57
67	Sweden	SWE	29
...	...	...	...
32	Iceland	ISL	1
47	Madagascar	MDG	1
34	Indonesia	IDN	1
36	Iraq	IRQ	1
78	Zimbabwe	ZWE	1

79 rows × 3 columns

In [33]:

```
world_map = px.choropleth(df_countries,
                          locations='ISO',
                          color='prize',
                          hover_name='birth_country_current',
                          color_continuous_scale=px.colors.sequential.matter)

world_map.update_layout(coloraxis_showscale=True,)

world_map.show()
```



In Which Categories are the Different Countries Winning Prizes?

In [34]:

```
cat_country = df_data.groupby(['birth_country_current', 'category'],
                              as_index=False).agg({'prize': pd.Series.count})
cat_country.sort_values(by='prize', ascending=False, inplace=True)
cat_country
```

Out[34]:

	birth_country_current	category	prize
204	United States of America	Medicine	78
206	United States of America	Physics	70
201	United States of America	Chemistry	55
202	United States of America	Economics	49
198	United Kingdom	Medicine	28
...	...	...	...
97	Iraq	Peace	1
99	Ireland	Medicine	1
100	Ireland	Physics	1
102	Israel	Economics	1
210	Zimbabwe	Peace	1

211 rows × 3 columns

In [35]:

```
merged_df = pd.merge(cat_country, top20_countries, on='birth_country_current')
# change column names
merged_df.columns = ['birth_country_current', 'category', 'cat_prize', 'total_prize']
merged_df.sort_values(by='total_prize', inplace=True)
merged_df
```

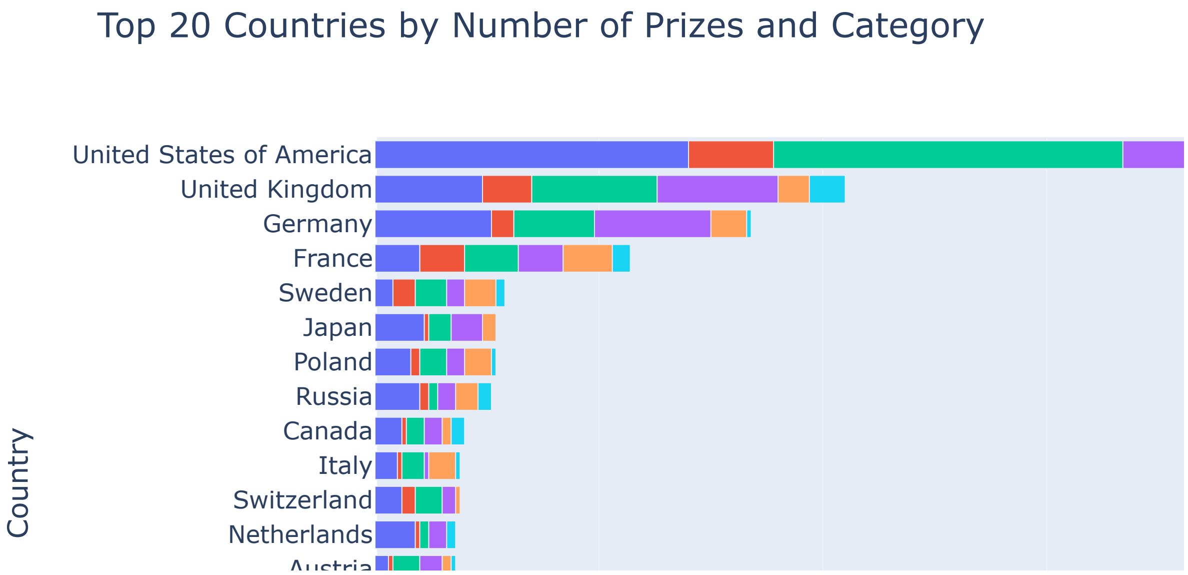
Out[35]:

	birth_country_current	category	cat_prize	total_prize
109	India	Physics	1	9
108	India	Peace	1	9
88	Belgium	Peace	3	9
89	Belgium	Medicine	3	9
90	Belgium	Chemistry	1	9
...	...	...	...	...
4	United States of America	Peace	19	281
3	United States of America	Economics	49	281
2	United States of America	Chemistry	55	281
1	United States of America	Physics	70	281
0	United States of America	Medicine	78	281

110 rows × 4 columns

In [36]:

```
cat_cntry_bar = px.bar(x=merged_df.cat_prize,
                       y=merged_df.birth_country_current,
                       color=merged_df.category,
                       orientation='h',
                       title='Top 20 Countries by Number of Prizes and Category')

cat_cntry_bar.update_layout(xaxis_title='Number of Prizes',
                            yaxis_title='Country')
cat_cntry_bar.show()
```



Number of Prizes Won by Each Country Over Time

In [37]:

```
prize_by_year = df_data.groupby(by=['birth_country_current', 'year'], as_index=False).count
prize_by_year = prize_by_year.sort_values('year')[['year', 'birth_country_current', 'prize'
prize_by_year
```

Out[37]:

	year	birth_country_current	prize
118	1901	France	2
346	1901	Poland	1
159	1901	Germany	1
312	1901	Netherlands	1
440	1901	Switzerland	1
...	...	...	...
31	2019	Austria	1
221	2020	Germany	1
622	2020	United States of America	7
533	2020	United Kingdom	2
158	2020	France	1

627 rows × 3 columns

In [38]:

```
cumulative_prizes = prize_by_year.groupby(by=['birth_country_current',
                                              'year']).sum().groupby(level=[0]).cumsum()
cumulative_prizes.reset_index(inplace=True)
cumulative_prizes
```

Out[38]:

	birth_country_current	year	prize
0	Algeria	1957	1
1	Algeria	1997	2
2	Argentina	1936	1
3	Argentina	1947	2
4	Argentina	1980	3
...	...	...	...
622	United States of America	2020	281
623	Venezuela	1980	1
624	Vietnam	1973	1
625	Yemen	2011	1
626	Zimbabwe	1960	1

627 rows × 3 columns

In [39]:

```
l_chart = px.line(cumulative_prizes,
                  x='year',
                  y='prize',
                  color='birth_country_current',
                  hover_name='birth_country_current')

l_chart.update_layout(xaxis_title='Year',
                      yaxis_title='Number of Prizes')

l_chart.show()
```

250
200
150
100
Number of Prizes

What are the Top Research Organisations?

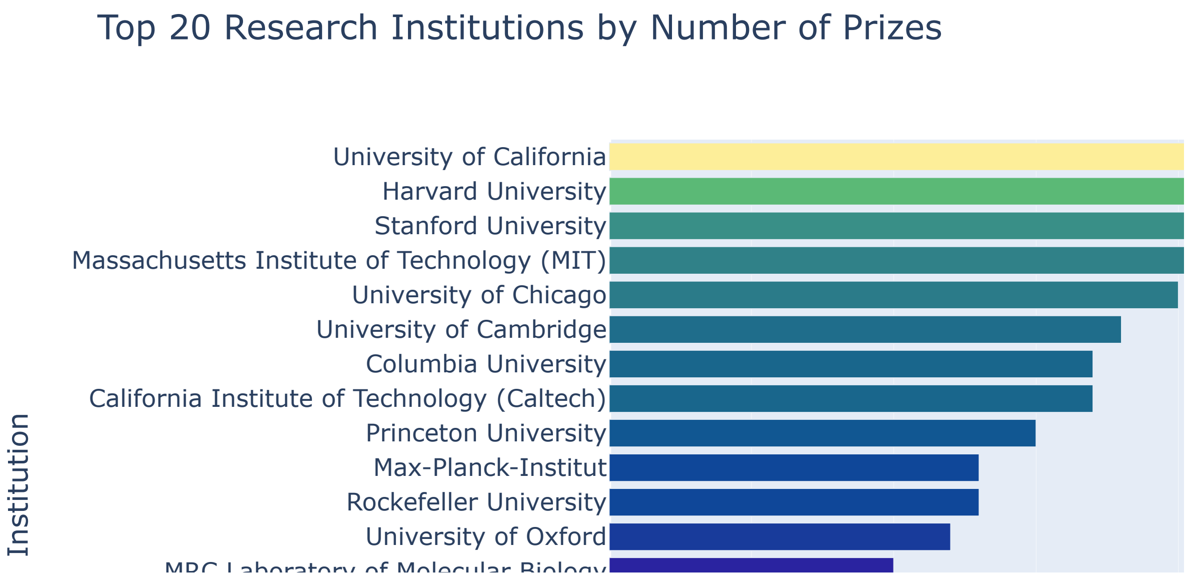
- Which organisations make up the top 20?
- How many Nobel prize winners are affiliated with the University of Chicago and Harvard University?

In [40]:

```
top20_orgs = df_data.organization_name.value_counts()[:20]
top20_orgs.sort_values(ascending=True, inplace=True)
```

In [41]:

```
org_bar = px.bar(x = top20_orgs.values,
                 y = top20_orgs.index,
                 orientation='h',
                 color=top20_orgs.values,
                 color_continuous_scale=px.colors.sequential.haline,
                 title='Top 20 Research Institutions by Number of Prizes')

org_bar.update_layout(xaxis_title='Number of Prizes',
                      yaxis_title='Institution',
                      coloraxis_showscale=False)
org_bar.show()
```

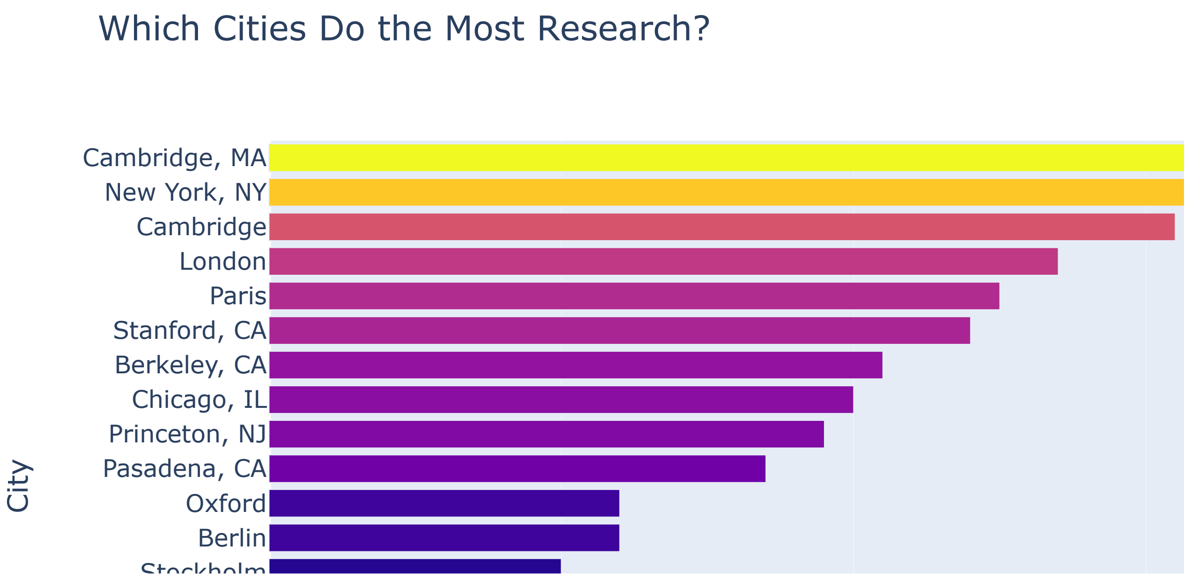


Which Cities Make the Most Discoveries?

Where do major discoveries take place?

In [42]:

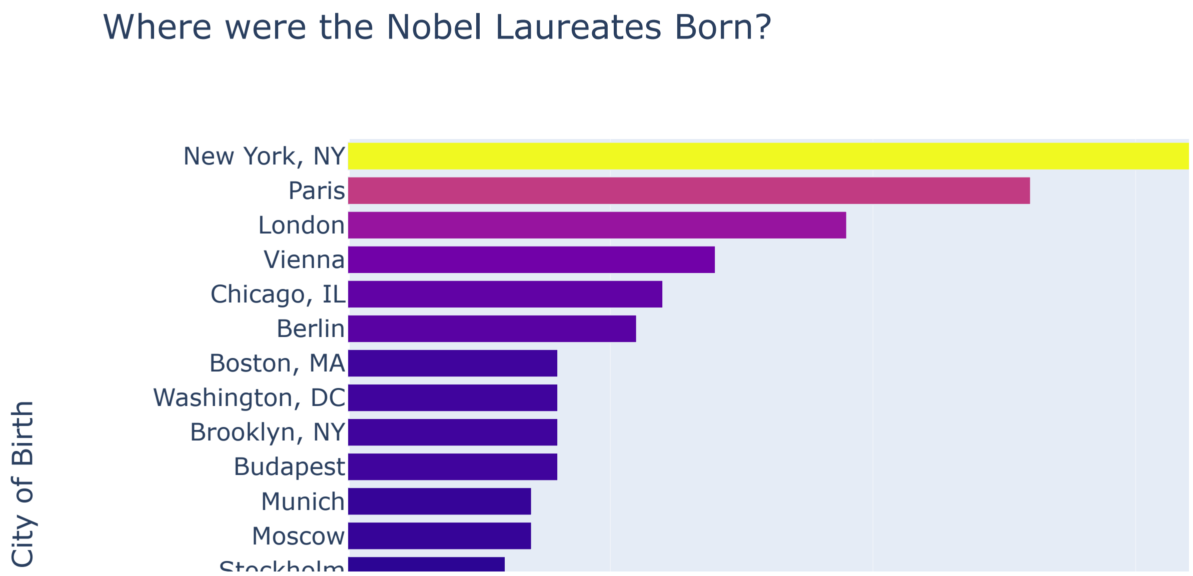
```
top20_org_cities = df_data.organization_city.value_counts()[:20]
top20_org_cities.sort_values(ascending=True, inplace=True)
city_bar2 = px.bar(x = top20_org_cities.values,
                   y = top20_org_cities.index,
                   orientation='h',
                   color=top20_org_cities.values,
                   color_continuous_scale=px.colors.sequential.Plasma,
                   title='Which Cities Do the Most Research?')

city_bar2.update_layout(xaxis_title='Number of Prizes',
                        yaxis_title='City',
                        coloraxis_showscale=False)
city_bar2.show()
```



Where are Nobel Laureates Born?

In [43]:

```
top20_cities = df_data.birth_city.value_counts()[:20]
top20_cities.sort_values(ascending=True, inplace=True)
city_bar = px.bar(x=top20_cities.values,
                  y=top20_cities.index,
                  orientation='h',
                  color=top20_cities.values,
                  color_continuous_scale=px.colors.sequential.Plasma,
                  title='Where were the Nobel Laureates Born?')

city_bar.update_layout(xaxis_title='Number of Prizes',
                       yaxis_title='City of Birth',
                       coloraxis_showscale=False)
city_bar.show()
```



Plotly Sunburst Chart: Combining Country, City, and Organisation

In [44]:

```
country_city_org = df_data.groupby(by=['organization_country',
                                       'organization_city',
                                       'organization_name'], as_index=False).agg({'prize':

country_city_org = country_city_org.sort_values('prize', ascending=False)
country_city_org
```

Out[44]:

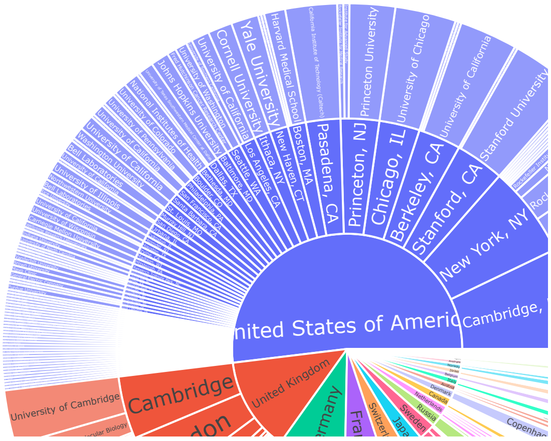
	organization_country	organization_city	organization_name	prize
205	United States of America	Cambridge, MA	Harvard University	29
280	United States of America	Stanford, CA	Stanford University	23
206	United States of America	Cambridge, MA	Massachusetts Institute of Technology (MIT)	21
209	United States of America	Chicago, IL	University of Chicago	20
195	United States of America	Berkeley, CA	University of California	19
...	...	...	...	...
110	Japan	Sapporo	Hokkaido University	1
111	Japan	Tokyo	Asahi Kasei Corporation	1
112	Japan	Tokyo	Kitasato University	1
113	Japan	Tokyo	Tokyo Institute of Technology	1
290	United States of America	Yorktown Heights, NY	IBM Thomas J. Watson Research Center	1

291 rows × 4 columns

In [45]:

```
burst = px.sunburst(country_city_org,
                    path=['organization_country', 'organization_city', 'organization_name'],
                    values='prize',
                    title='Where do Discoveries Take Place?',
                   )

burst.update_layout(xaxis_title='Number of Prizes',
                    yaxis_title='City',
                    coloraxis_showscale=False)

burst.show()
```

Where do Discoveries Take Place?


nited States of Americ
Cambridge,
New York, NY
Stanford, CA
Berkeley, CA
Chicago, IL
Princeton, NJ
Pasadena, CA
Boston, MA
New Haven, CT
Ithaca, NY
Los Angeles, CA
Seattle, WA
Stanford University
University of California
University of Chicago
Princeton University
California Institute of Technology (Caltech)
Harvard Medical School
Yale University
Cornell University
University of California
Johns Hopkins University
National Institutes of Health
United Kingdom
Cambridge
don
University of Cambridge
rmany
Fra
Switzer
Japa
Sweden
Russia
Netherlands
Canada
Denmark
Copenhag