

# Assignment 4

```
In [1]: # load packages
import pandas as pd
import altair as alt
from vega_datasets import data
```

```
In [2]: # load data
countries = pd.read_csv('countries.csv')
flunet = pd.read_csv('flunet2010_11countries_106.csv')
```

```
In [3]: countries.head()
```

Out[3]:

	id	name	alpha2	alpha3
0	4	Afghanistan	af	afg
1	8	Albania	al	alb
2	12	Algeria	dz	dza
3	20	Andorra	ad	and
4	24	Angola	ao	ago

```
In [4]: flunet.head()
```

```
Out[4]:
```

	week	Afghanistan	Argentina	Australia	Canada	China	Colombia	Egypt	Germany	Ireland	South Africa	USA
0	1	5.0	4	2	41	2179	36	739	26.0	23	0	366
1	2	13.0	21	1	15	2213	36	396	24.0	8	1	396
2	3	4.0	6	1	8	2228	14	192	18.0	4	0	447
3	4	0.0	1	0	14	2027	11	80	NaN	8	0	402
4	5	0.0	4	1	12	1813	8	56	NaN	4	0	404

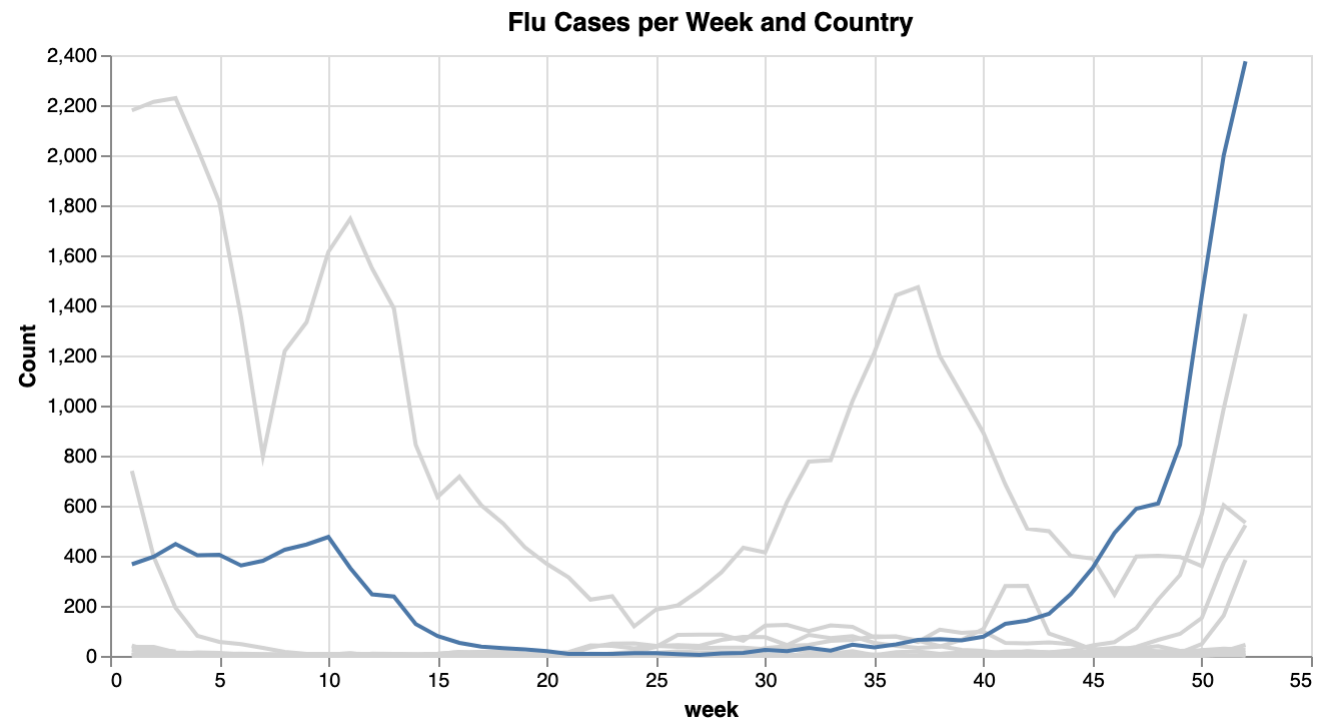
## Q1 Line Chart with Selection

```
In [5]: # preprocess: convert wide form into long form
country_names = flunet.columns[1:].to_numpy()
flunet_long = flunet.melt('week', var_name='Country', value_name='Count')

# set color and selection properties
input_dropdown = alt.binding_select(options=country_names)
selection = alt.selection_single(fields=['Country'], bind=input_dropdown, name='Select ')
color = alt.condition(selection,
                       alt.Color('Country:N', legend=None), # too much colors, hide legends
                       alt.value('lightgray'))

# make the plot
alt.Chart(flunet_long).mark_line().encode(
    x='week',
    y='Count',
    color=color,
    tooltip=['Country', 'week', 'Count']
).add_selection(
    selection
).properties(
    title='Flu Cases per Week and Country',
    width=600
)
```

Out[5]:



Select\_\_Country  ▼

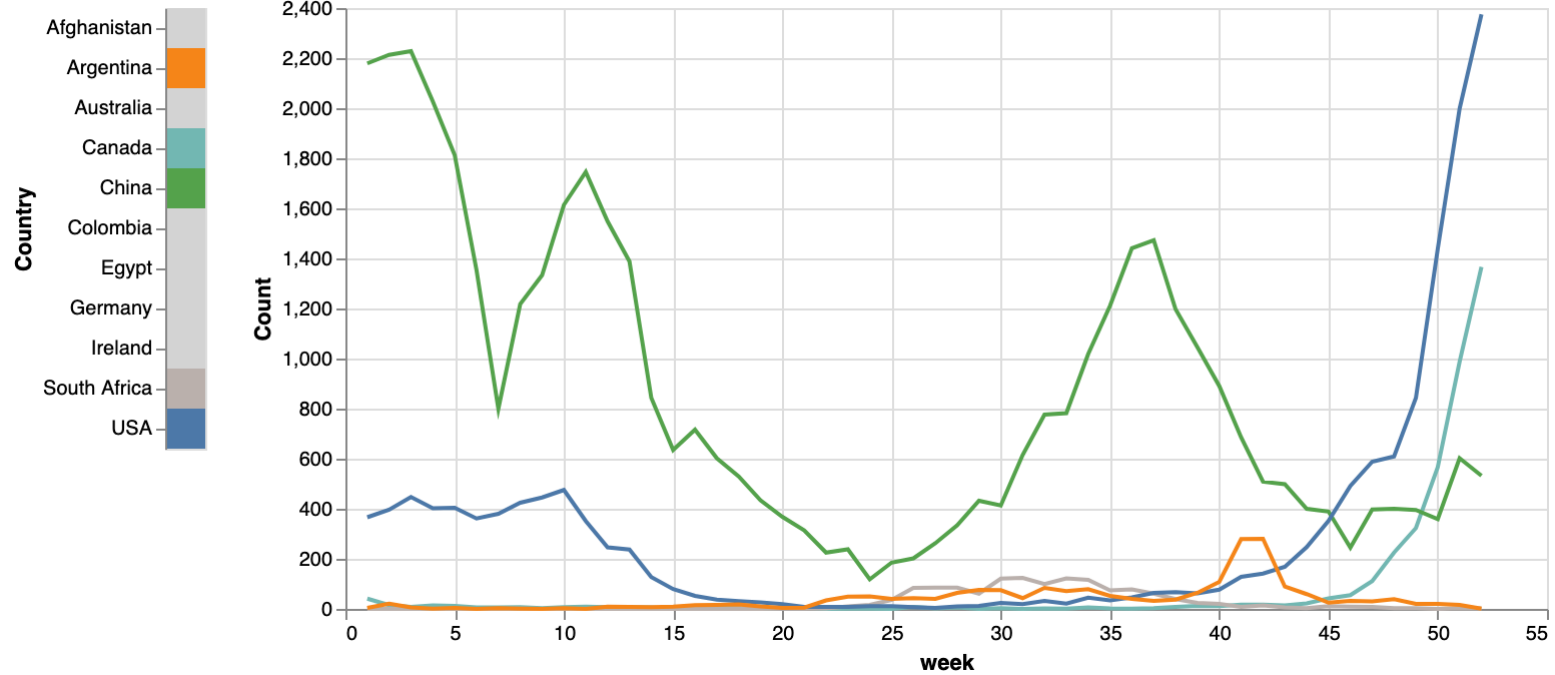
```
In [6]: # bonus
# add selection
selection = alt.selection_multi(fields=['Country'])
color = alt.condition(selection, alt.Color('Country:N', legend=None), alt.value('lightgray'))

# make a checkbox
make_selector = alt.Chart(
    pd.DataFrame({'Country': country_names})
).mark_rect().encode(y='Country', color=color).add_selection(selection)

# make the line chart
line_chart = alt.Chart(
    flunet_long
).mark_line().encode(
    x='week',
    y='Count',
    color=color,
    tooltip=['Country', 'week', 'Count']
).transform_filter(selection).properties(
    title='Flu Cases per Week and Country',
    width=600
)

# combine plots
make_selector | line_chart
# shift click to make multiple selection!!!
```

Out[6]:



Q2 Choropleth

```
In [7]: # gather data, groupby and merge datasets
world = alt.topo_feature(data.world_110m.url, 'countries')
source = flunet_long.groupby('Country')['Count'].sum().reset_index().merge(
    countries, left_on='Country', right_on='name', how='outer'
)

# make a background to indicate missing values
background = alt.Chart(world).mark_geoshape(
    fill='lightgray',
    stroke='white'
).project(
    "equirectangular"
).properties(
    width=500,
    height=300
)

# make a foreground to indicate present values
foreground = alt.Chart(world).mark_geoshape().encode(
    color='Count:Q',
    tooltip=['name:N', 'Count:Q']
).transform_lookup(
    lookup='id',
    from_=alt.LookupData(source, 'id', ['name', 'Count'])
).project(
    type='equirectangular'
).properties(
    width=500,
    height=300,
    title='Annual Flu Cases per Country'
)

# combine plots
background + foreground
```

Out[7]:

