Assignment 3

Author: Sheng Yang

In [1]:
```
# import packages
import pandas as pd
import altair as alt
```

In [2]:
```
# load data
fatality = pd.read_csv('ukDriverFatalities.csv')
fatality.head()
```

Out[2]:

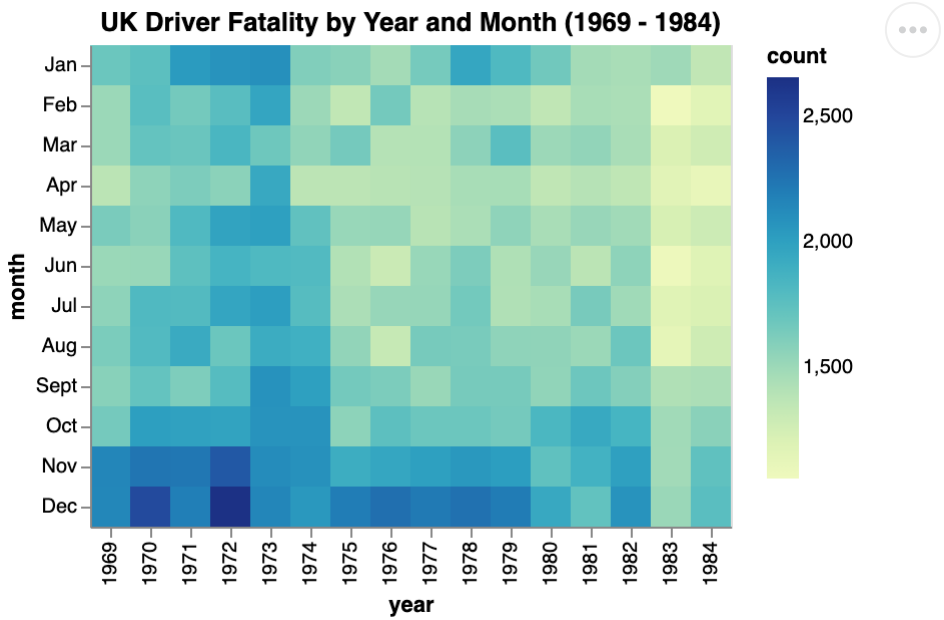
	month	year	count
0	0	1969	1687
1	1	1969	1508
2	2	1969	1507
3	3	1969	1385
4	4	1969	1632

Q1. Heatmap

In [3]:
```
# convert numbers the correct month (1 ~ 12)
fatality_copy = fatality.copy()
months = ['Jan', 'Feb', 'Mar', 'Apr', 'May', 'Jun', 'Jul', 'Aug', 'Sept', 'Oct', 'Nov', 'Dec']
month_dict = dict(zip(range(12), months))
fatality_copy.month = fatality.month.replace(month_dict)

# heatmap
alt.Chart(fatality_copy).mark_rect().encode(
    x='year:O',
    y=alt.Y('month:O', sort=months),
    color='count:Q',
    tooltip=['year', 'month', 'count']
).properties(
    title='UK Driver Fatality by Year and Month (1969 - 1984)'
)
```

Out[3]:



Q2. Line Chart

In [4]:
```
# line chart
# manually adjust domain to make line chart a 45 degree tilt
alt.Chart(fatality).mark_line().encode(
    x=alt.X('year:O', axis=alt.Axis(labelAngle=0)),
    y=alt.Y('sum(count)',
            title='fatality',
            scale=alt.Scale(domain=(12000, 26000))),
    tooltip=['year:O', alt.Tooltip('sum(count)', title='fatality')]
).properties(
    title='UK Driver Death Toll by Year (1969 - 1984)',
    width=500
)
```

Out[4]: