```python
In [1]:  %matplotlib inline
         import pandas as pd
         import matplotlib.pyplot as plt
         import seaborn as sns
         sns.set(style="darkgrid")
```

```python
In [2]:  df = pd.read_csv('fortune500.csv')
         df.head()
```

Out[2]:

	Year	Rank	Company	Revenue (in millions)	Profit (in millions)
0	1955	1	General Motors	9823.5	806
1	1955	2	Exxon Mobil	5661.4	584.8
2	1955	3	U.S. Steel	3250.4	195.4
3	1955	4	General Electric	2959.1	212.6
4	1955	5	Esmark	2510.8	19.1

```python
In [3]:  df.tail()
         df.columns = ['year', 'rank', 'company', 'revenue', 'profit']

         non_numberic_profits = df.profit.str.contains('[^0-9.-]')
```
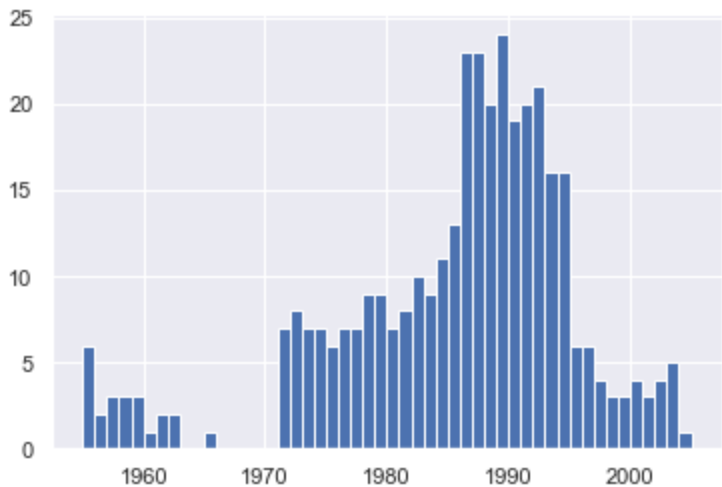
```python
In [4]:  df.loc[non_numberic_profits].head()
```

Out[4]:

|	year	rank	company	revenue	profit
228	1955	229	Norton	135.0	N.A.
290	1955	291	Schlitz Brewing	100.0	N.A.
294	1955	295	Pacific Vegetable Oil	97.9	N.A.
296	1955	297	Liebmann Breweries	96.0	N.A.
352	1955	353	Minneapolis-Moline	77.4	N.A.

```python
In [5]:  bin_sizes, _, _ = plt.hist(df.year[non_numberic_profits], bins=range(1955, 2006))
```



```python
In [6]:  df = df.loc[~non_numberic_profits]
         df.profit = df.profit.apply(pd.to_numeric)
```

```python
In [7]:  group_by_year = df.loc[:, ['year', 'revenue', 'profit']].groupby('year')
         avgs = group_by_year.mean()
         x = avgs.index
         y1 = avgs.profit
         def plot(x, y, ax, title, y_label):
             ax.set_title(title)
             ax.set_ylabel(y_label)
             ax.plot(x, y)
             ax.margins(x=0, y=0)
```

```python
In [8]:  fig, ax = plt.subplots()
         plot(x, y1, ax, 'Increase in mean Fortune 500 company profits from 1955 to 2005', 'Pro
```
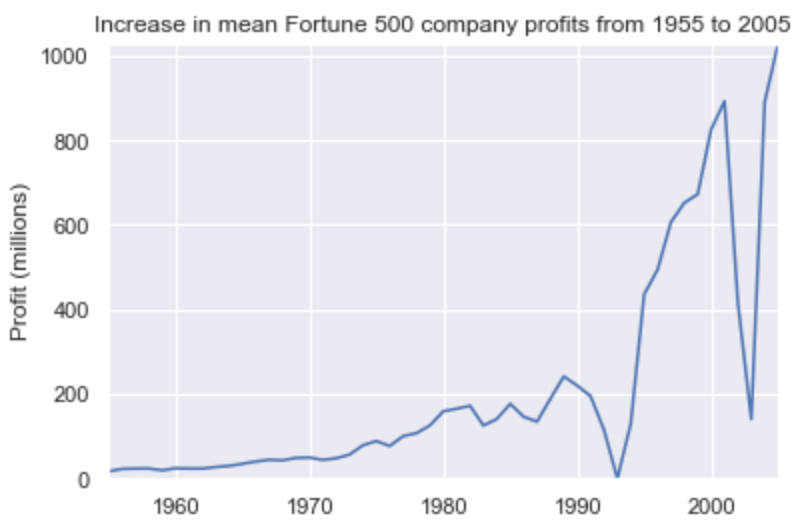


```python
In [9]:  y2 = avgs.revenue
         fig, ax = plt.subplots()
         plot(x, y2, ax, 'Increase in mean Fortune 500 company revenues from 1955 to 2005', 'Re

         def plot_with_std(x, y, stds, ax, title, y_label):
             ax.fill_between(x, y - stds, y + stds, alpha=0.2)
             plot(x, y, ax, title, y_label)
         fig, (ax1, ax2) = plt.subplots(ncols=2)
         title = 'Increase in mean and std Fortune 500 company %s from 1955 to 2005'
         stds1 = group_by_year.std().profit.values
         stds2 = group_by_year.std().revenue.values
         plot_with_std(x, y1.values, stds1, ax1, title % 'profits', 'Profit (millions)')
         plot_with_std(x, y2.values, stds2, ax2, title % 'revenues', 'Revenue (millions)')
         fig.set_size_inches(14, 4)
         fig.tight_layout()
```