Welcome to your third case study on Data Visualization

- In this task you will asked to draw few plots on the weather data you scraped earlier.
- The data is provided in the data.csv file in the current working directory.
- You will be specifically using seaborn package to draw the plots.
- Your plots must match with the expected plots provided in each task.

For each of the plot assign the plot object to variables mentioned in the comments.

```
- for example if your drawing a scatter plot

plot_var = seaborn.scatter()
```

Run the below cell to import the necessary packages.

In [1]:
```
import pandas as pd
import seaborn as sns
import matplotlib.pyplot as plt
import warnings
warnings.filterwarnings('ignore')
```

In [4]:
```
###Read the dataset form data.csv file
df = pd.read_csv('data.csv')
df['Day'] = pd.to_datetime(df['Day'], format='%d/%m/%Y')
df.head()
```

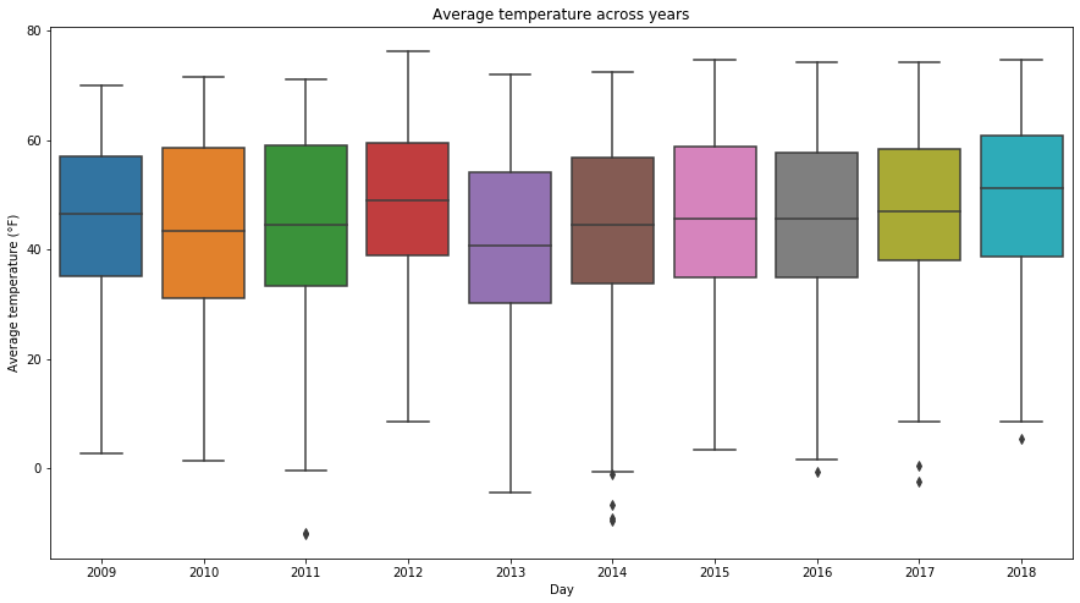
Out[4]:

	Day	Average temperature (°F)	Average humidity (%)	Average dewpoint (°F)	Average barometer (in)	Average windspeed (mph)	Average gustspeed (mph)	Averag directio (°de
0	2009-01-01	37.8	35	12.7	29.7	26.4	36.8	2
1	2009-01-02	43.2	32	14.7	29.5	12.8	18.0	2
2	2009-01-03	25.7	60	12.7	29.7	8.3	12.2	2
3	2009-01-04	9.3	67	0.1	30.4	2.9	4.5	
4	2009-01-05	23.5	30	-5.3	29.9	16.7	23.1	2

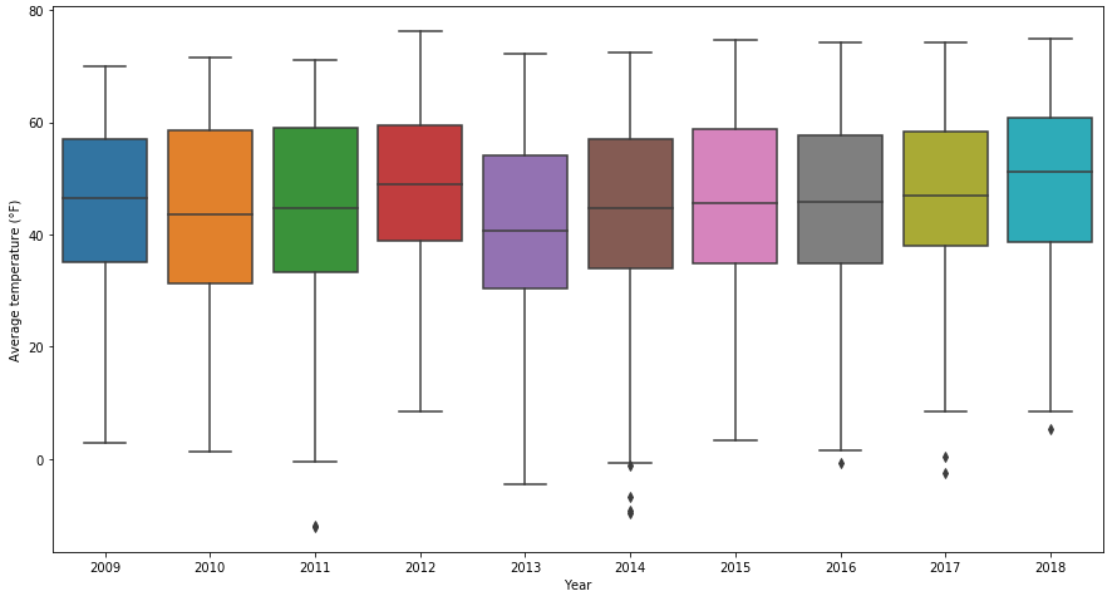
Draw a box plot on the average_temperature column across each year.

- use seaborn.boxplot()
- set the height and width to 8 and 15 repectively.
- assign the plot object to variable plot1.

Expected plot



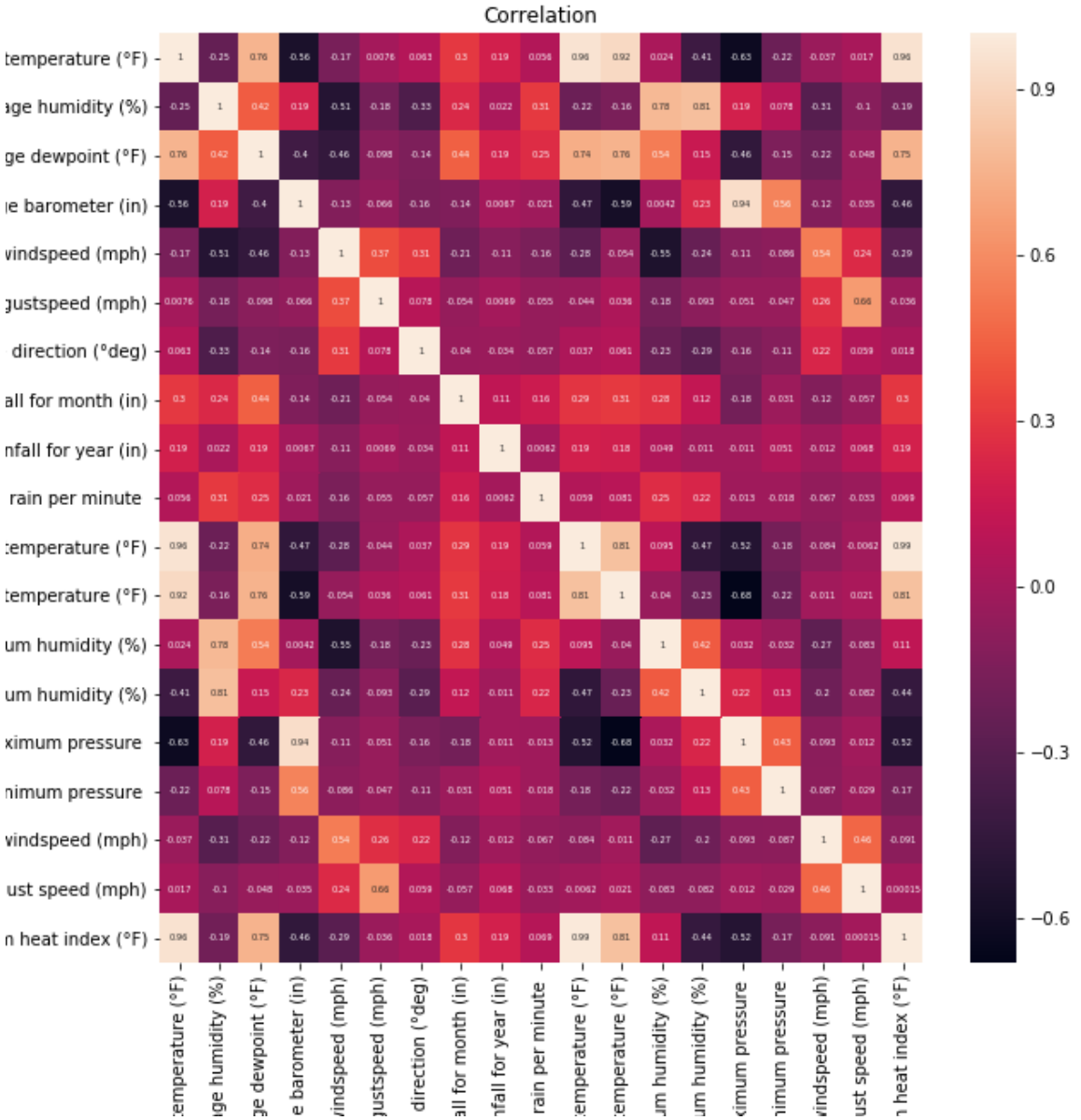
In [6]:
```
### Start code here

df_temp =df
df_temp['Year'] = df['Day'].dt.year

plt.subplots(figsize=(15,8))
plot1 = sns.boxplot(x='Year', y='Average temperature (°F)', data
=df_temp)
```



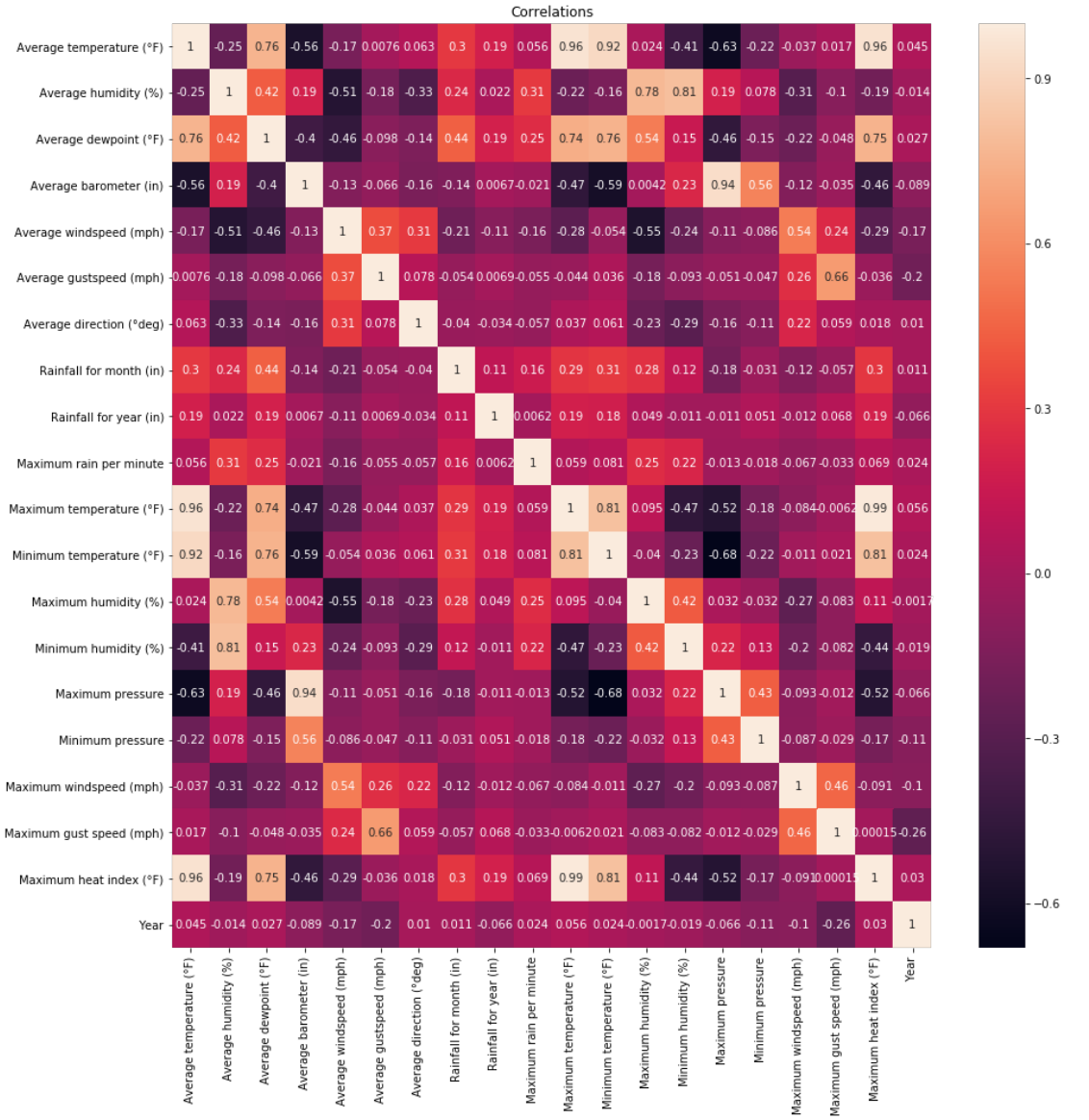
Draw correlation heatmap for all the available features.

```
- make sure the correlation values are annoted for each combination o
f features.
- assign the plot object to variable plot2
```

Expected plot



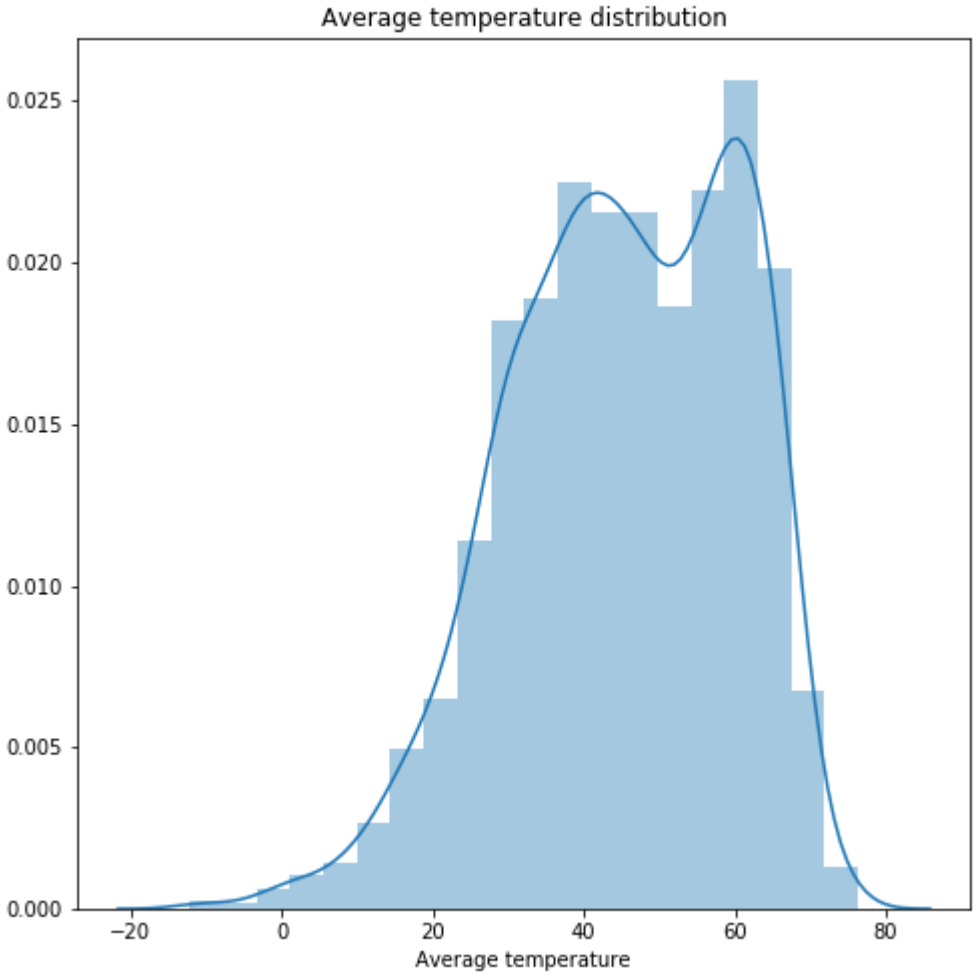
In [7]:
```
###Start code here

plt.subplots(figsize=(15,15))
plot2 = sns.heatmap(df.corr(), annot=True).set(title="Correlations")
```



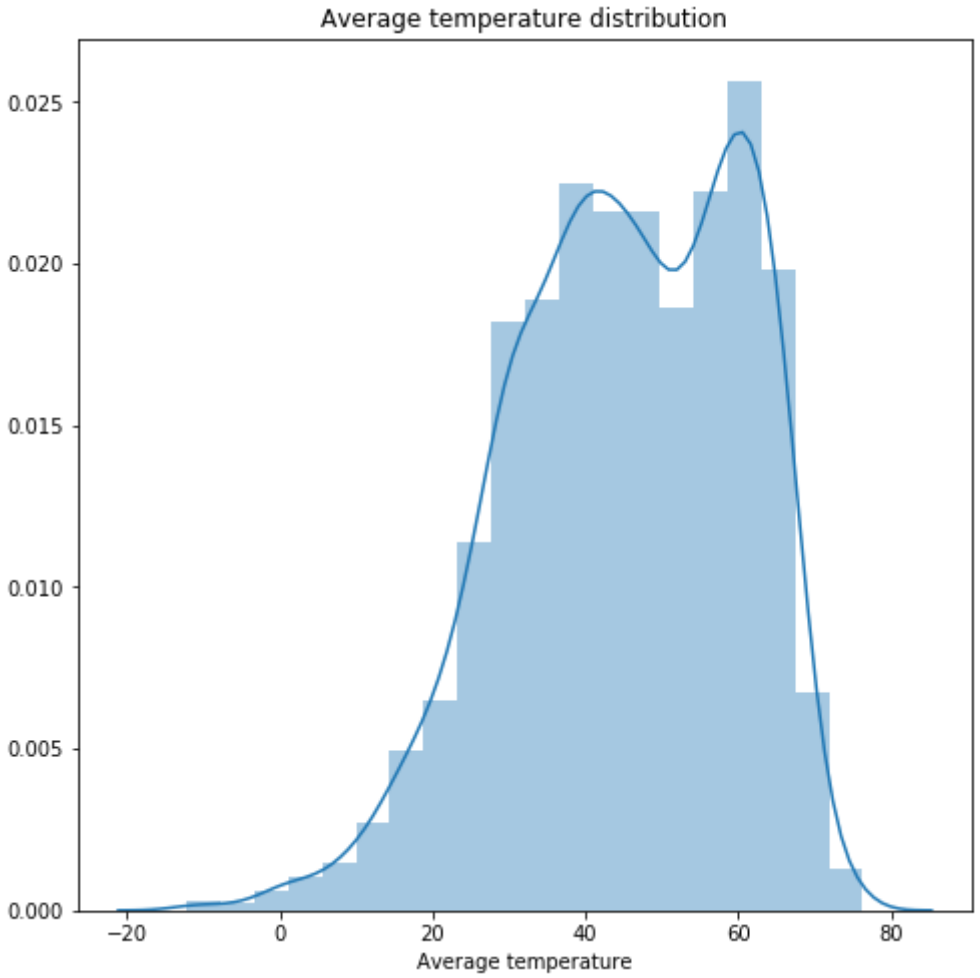
Draw a distribution plot using seaborn for average_temperature column.

- set bins to 20
- set the height and width to 8,8.
- assign the plot object to plot3 variable

Expected plot



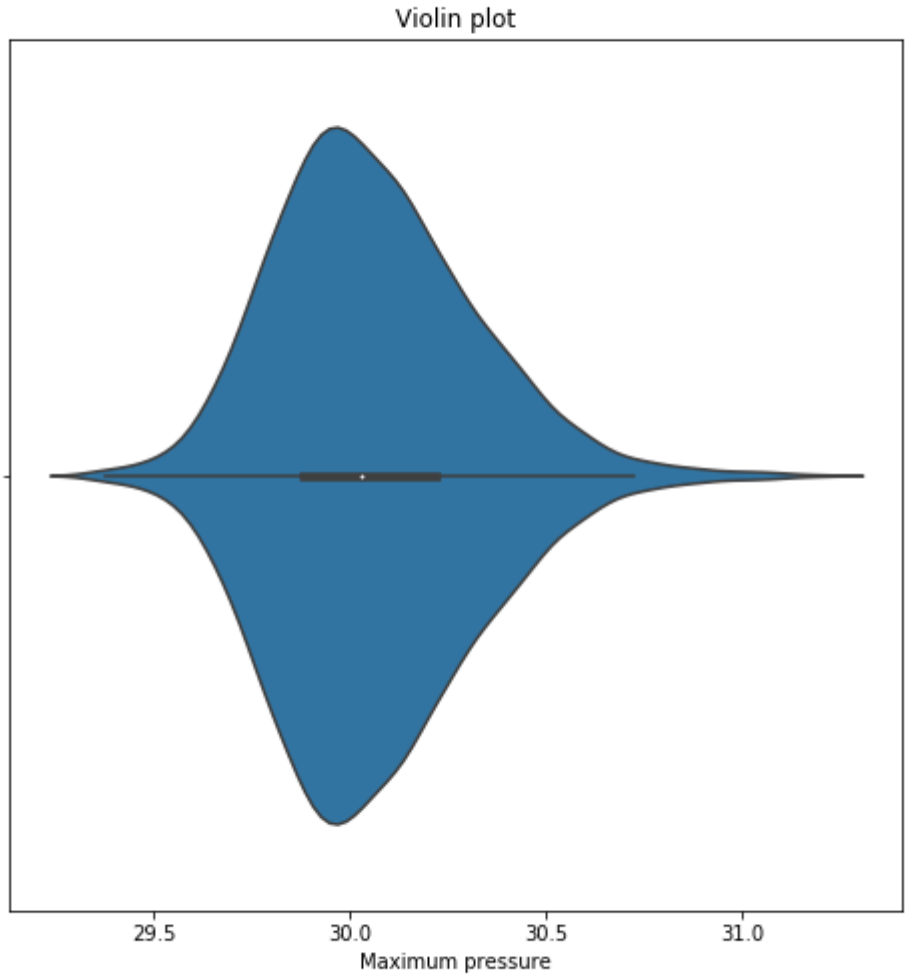
In [8]:
```
###Start code here


plt.subplots(figsize=(8, 8))
plot3 = sns.distplot(df['Average temperature (°F)'], bins=20, kde=True, axlabel="Average temperature").set(title="Average temperature distribution")
```



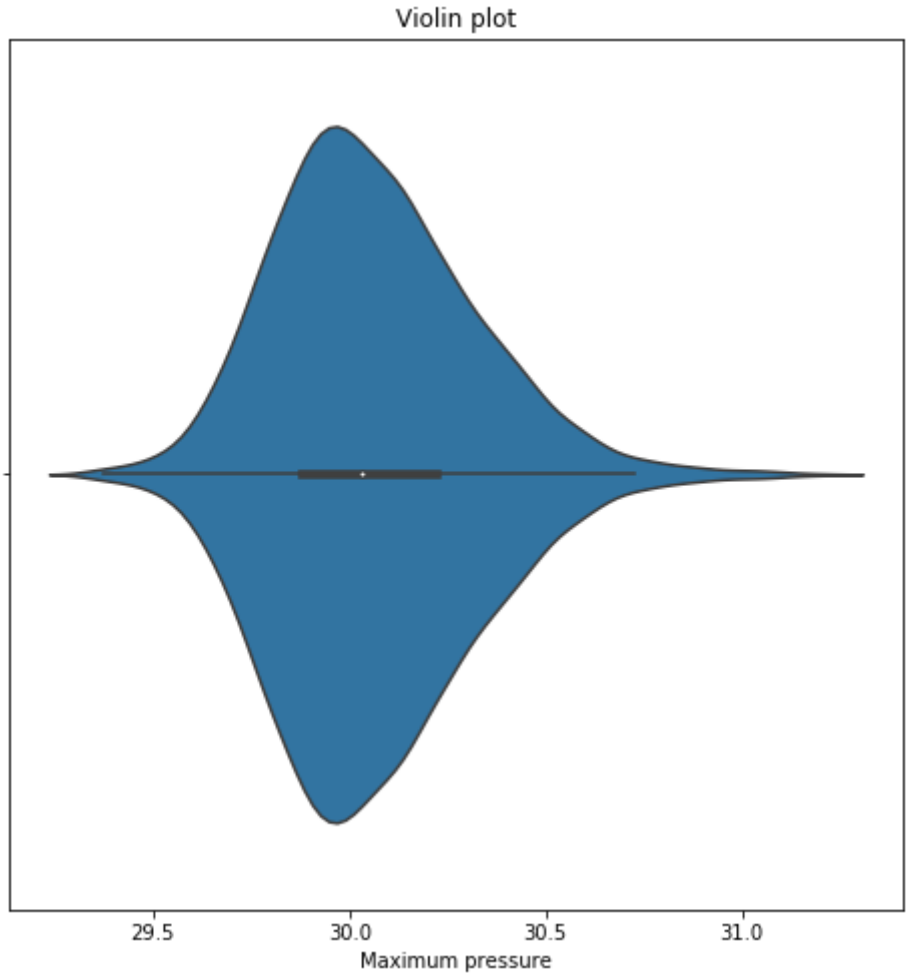
Draw seaborn violin plot on mximum_pressure column.

- set gridsize to 100
- set figsize to 8,8
- assign the plot object to variable plot4

Expected plot



In [9]:
```
###Start code here


plt.subplots(figsize=(8,8))
plot4 = sns.violinplot(x=df['Maximum pressure'], axlabbel='Maxim
um Pressure', gridsize=100).set(title="Violin plot")
```



Run the below cell to save your plot objects.

In [10]:
```
import pickle
with open("plot1.pickle", "wb") as file:
    pickle.dump(plot1, file)

with open("plot2.pickle", "wb") as file:
    pickle.dump(plot2, file)

with open("plot3.pickle", "wb") as file:
    pickle.dump(plot3, file)

with open("plot4.pickle", "wb") as file:
    pickle.dump(plot4, file)
```

In [ ]:
```

```