

```
In [1]: ▶ import numpy as np
import pandas as pd
import seaborn as sns
import matplotlib.pyplot as plt
```

Read file and display it. This is to get an idea of the data I am working with

```
In [4]: ▶ myDataFromFile = pd.read_csv("datasets/007ExerciseFile.csv")
display(myDataFromFile)
```

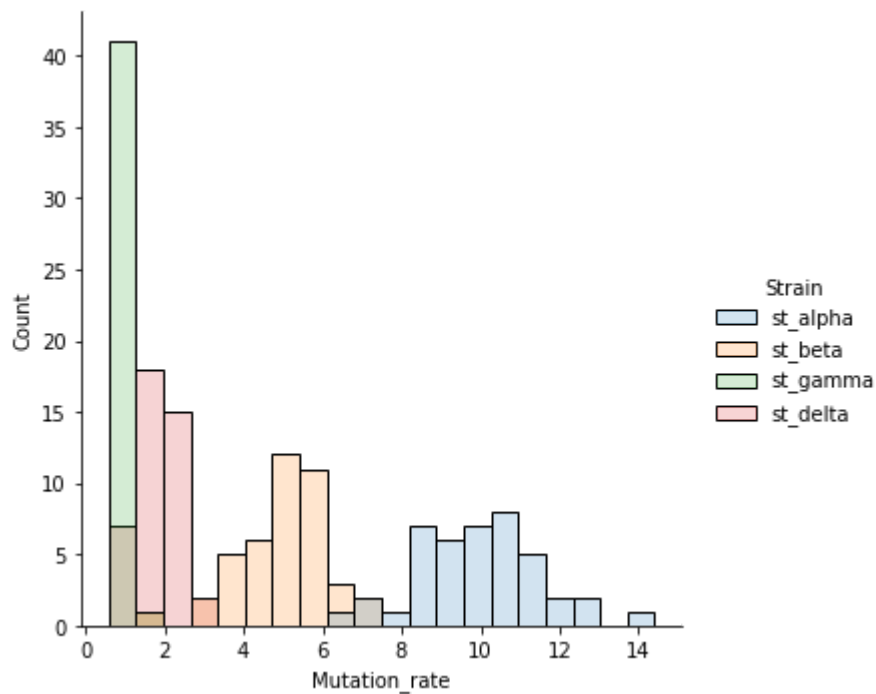
	Strain	Mutation_rate
0	st_alpha	10.612005
1	st_alpha	12.586371
2	st_alpha	8.997583
3	st_alpha	11.681775
4	st_alpha	14.408237
...
163	st_delta	2.716249
164	st_delta	2.467378
165	st_delta	2.119801
166	st_delta	1.316537
167	st_delta	2.060472

168 rows × 2 columns

Make an informative histogram of the mutation rates of the 4 strains

```
In [13]: sns.displot(myDataFromFile, x="Mutation_rate", hue="Strain", kind="hist", alp
```

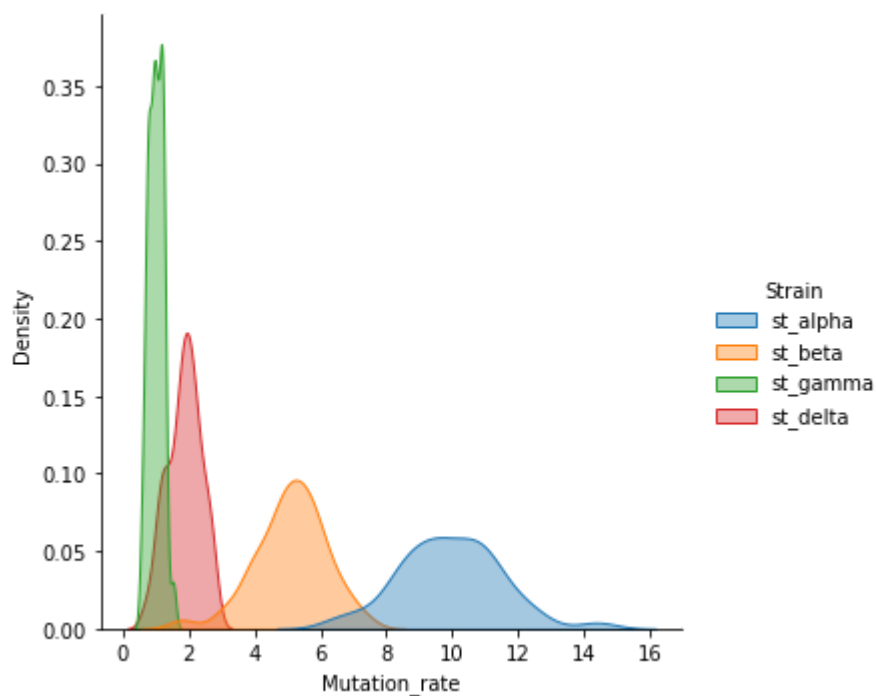
```
Out[13]: <seaborn.axisgrid.FacetGrid at 0x1dc248ad400>
```



Make an informative KDE plot

```
In [23]: sns.displot(myDataFromFile, x="Mutation_rate", hue="Strain", kind="kde", fill
```

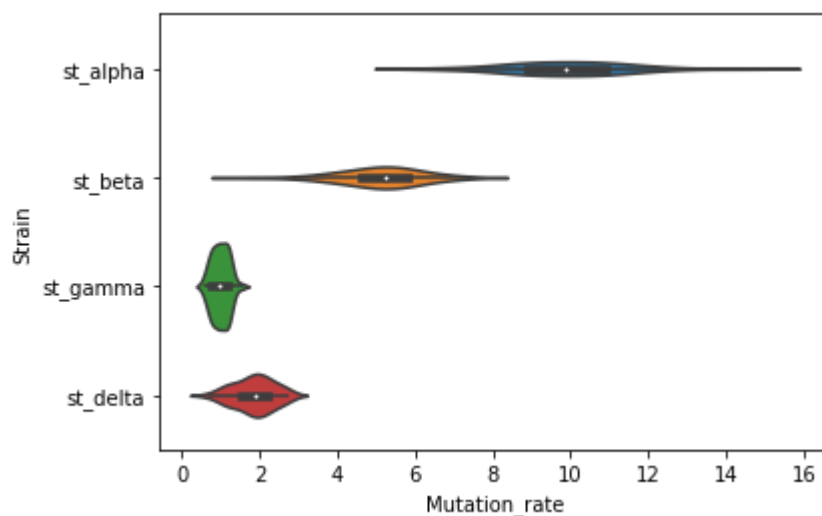
```
Out[23]: <seaborn.axisgrid.FacetGrid at 0x1dc249577c0>
```



Make a violin plot of the data

```
In [25]: sns.violinplot(data=myDataFromFile, x="Mutation_rate", y="Strain")
```

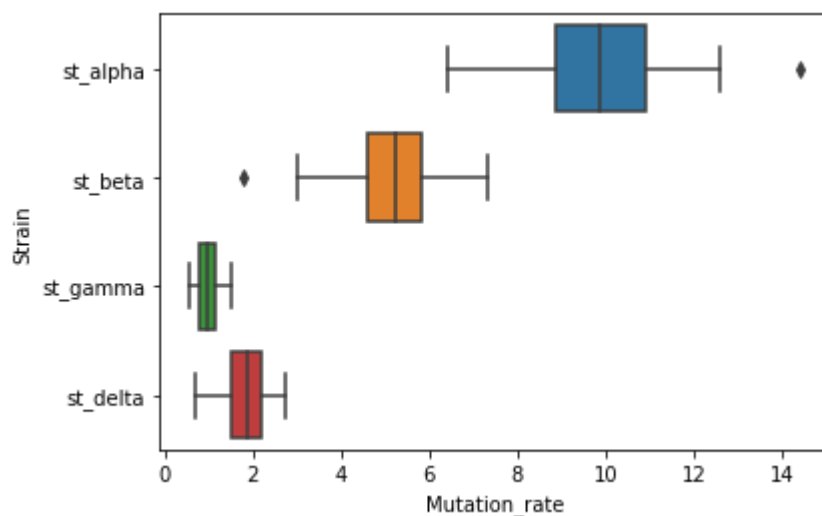
```
Out[25]: <AxesSubplot:xlabel='Mutation_rate', ylabel='Strain'>
```



Make a boxplot of the data

```
In [26]: sns.boxplot(data=myDataFromFile, x="Mutation_rate", y="Strain")
```

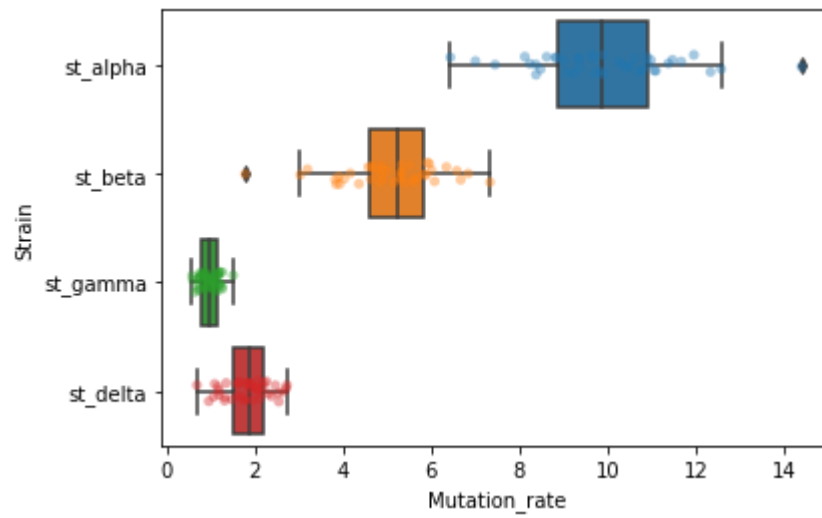
```
Out[26]: <AxesSubplot:xlabel='Mutation_rate', ylabel='Strain'>
```



Make a boxplot with an overlaid strip chart

```
In [27]: sns.boxplot(data=myDataFromFile, x="Mutation_rate", y="Strain")  
sns.stripplot(data=myDataFromFile, x="Mutation_rate", y="Strain", alpha=0.4)
```

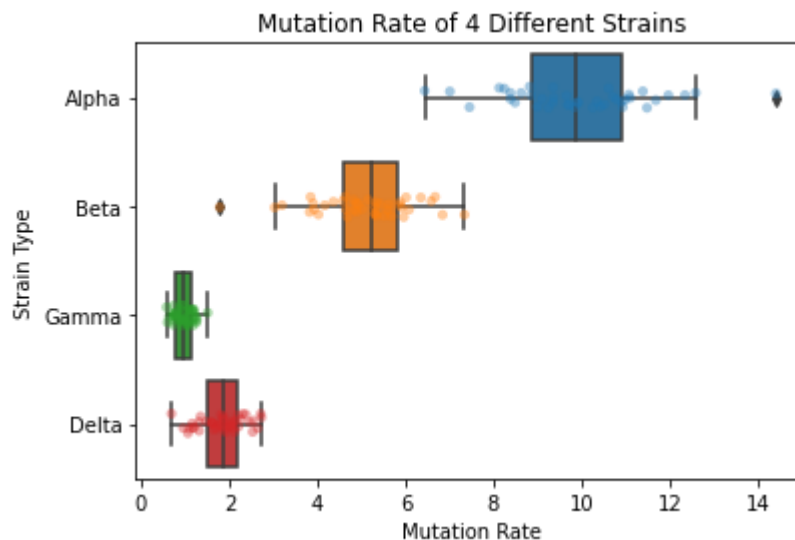
```
Out[27]: <AxesSubplot:xlabel='Mutation_rate', ylabel='Strain'>
```



Selected graph with English labels

```
In [56]: sns.boxplot(data=myDataFromFile, x="Mutation_rate", y="Strain")
sns.stripplot(data=myDataFromFile, x="Mutation_rate", y="Strain", alpha=0.4)
plt.title("Mutation Rate of 4 Different Strains")
plt.xlabel("Mutation Rate")
plt.ylabel("Strain Type")
plt.yticks([0, 1, 2, 3], ['Alpha', 'Beta', 'Gamma', 'Delta'])
```

```
Out[56]: ([<matplotlib.axis.YTick at 0x1dc27960df0>,
<matplotlib.axis.YTick at 0x1dc27960850>,
<matplotlib.axis.YTick at 0x1dc27a3cb50>,
<matplotlib.axis.YTick at 0x1dc27a7db50>],
[Text(0, 0, 'Alpha'),
Text(0, 1, 'Beta'),
Text(0, 2, 'Gamma'),
Text(0, 3, 'Delta')])
```



The graph demonstrates that each strain is mutating at a different rate with the Alpha strain mutating the fastest and the Gamma strain mutating the slowest.

In []: