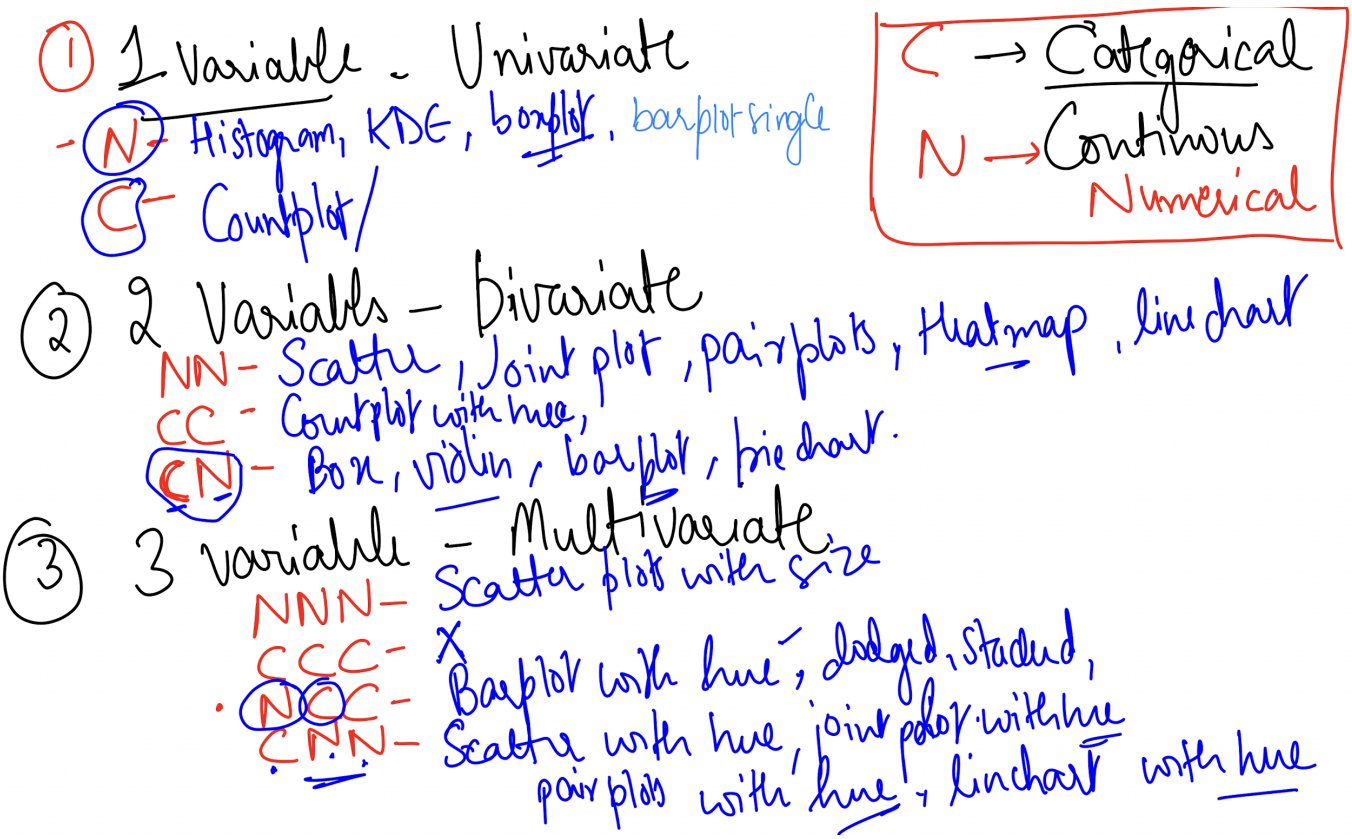
```
import numpy as np
import pandas as pd
import matplotlib.pyplot as plt
import seaborn as sns


iris=sns.load_dataset("iris")
tips=sns.load_dataset("tips")
```




```
iris
```

```
iris["species"].unique()
```

```
array(['setosa', 'versicolor', 'virginica'], dtype=object)
```

```
iris["species"].nunique()
```

```
3
```

```
iris["species"].value_counts()
```

```
setosa        50
versicolor    50
```

```
virginica     50
Name: species, dtype: int64
```

```
tips["smoker"].value_counts()
```

```
No     151
Yes     93
Name: smoker, dtype: int64
```

```
sns.countplot(data=iris,
              x="species")
```

```
sns.countplot(data=tips,
              x="smoker")
plt.show()
```

```
# petal_length and petal_width
```

```
sns.scatterplot(data=iris,
                x="petal_length",
                y="petal_width")
plt.show()
```

```
sns.scatterplot(data=iris,x="sepal_length",y="sepal_width",hue="species")
plt.show()
```

```
setosa=iris.loc[iris["species"]=="setosa"]
versicolor=iris.loc[iris["species"]=="versicolor"]
virginica=iris.loc[iris["species"]=="virginica"]
plt.scatter(x=setosa["petal_length"],y=setosa["petal_width"],c="r")
plt.scatter(x=versicolor["petal_length"],y=versicolor["petal_width"],c="g")
plt.scatter(x=virginica["petal_length"],y=virginica["petal_width"],c="b")
plt.xlabel("petal_length")
plt.ylabel("petal_width")
plt.legend(["setosa","versicolor","virginica"])

plt.show()
```

```
# 2 variable 2 plots in 1 plot
```

```
sns.jointplot(data=iris,
              x="petal_length",
              y="petal_width")
```

```
sns.jointplot(data=iris,
              x="petal_length",
              y="petal_width",
              hue="species")
```

```
sns.jointplot(data=iris,
              x="petal_length",
              y="petal_width",
              kind="reg",)
```

```
sns.jointplot(data=iris,
              x="petal_length",
              y="petal_width",
              kind="hist",
              hue="species")
```

```
sns.jointplot(data=iris,
              x="petal_length",
              y="petal_width",
              kind="hex")
```

```
iris
```

```
sns.pairplot(data=iris)
```

```
sns.pairplot(data=iris,hue="species")
```

```
# c ="sepal_length"
```

```
# n= "species"

sns.boxplot(data=iris,
            x="species",
            y="sepal_length")
```

```
# data=np.array([2,20,22,42])

# np.median(data)

# sns.boxplot(data)
plt.show()
```

```
data=np.array([-10,3,4,5,6,7,9,40])

sns.boxplot(x=data)
```

```
np.median(data)
```

```
5.5
```

```
sns.violinplot(data=iris,
            x="species",
            y="sepal_length")
```

```
iris.corr()
```

```
sns.heatmap(iris.corr(),annot=True)
```

```
sns.lineplot(data=iris,
             x="sepal_length",
             y="sepal_width",
             hue="species")
```

```
tips
```

```
sns.countplot(data=tips,
              x="sex")
```

```
sns.countplot(data=tips,
              x="sex",
             hue="smoker")
```

```
sns.countplot(data=tips,
              x="smoker",
             hue="sex")
```

```
sns.countplot(data=tips,
              x="smoker",
             hue="sex",
              dodge=False)
```

```
sns.barplot(data=tips,
           x="day",
           y="total_bill")
```

```
sns.barplot(data=tips,
           x="day",
```

```
            y="total_bill",
            estimator=np.min)
```

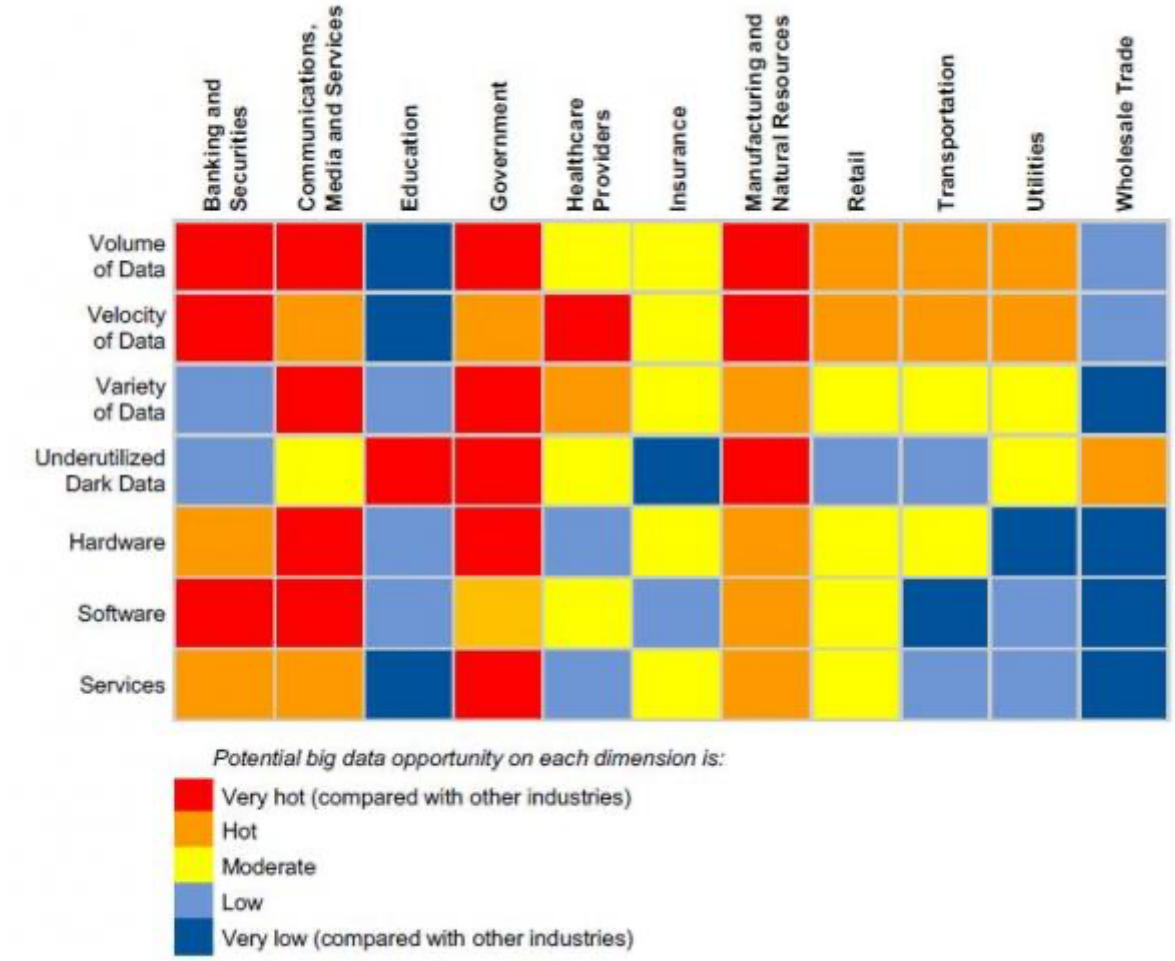
```
sns.barplot(data=tips,
            x="day",
            y="total_bill",
            estimator=np.max)
```

```
sns.barplot(data=tips,
            x="day",
            y="total_bill",
            hue="sex")
```

```
sns.barplot(data=tips,
            x="day",
            y="total_bill",
            hue="smoker")
```

```
sns.scatterplot(data=tips,
                x="tip",
                y="total_bill",
                hue="sex",
                 size="size"
                )
```

Figure 2. Big Data Opportunity Heat Map by Industry



Source: Gartner (July 2012)

Colab paid products - Cancel contracts here