

3.Presidents_Datasets_jithin

August 6, 2018

1 Pandas Basics - Data Analysis

Summary - In this Exploratory Data Analysis Exercise, we are going to analyse presidents heights dataset.

In this dataset there are very less observation and one feature hence, it is easy to understand.

- 1) Importing libraries / dataset.
- 2) Statistical Analysis
- 3) Sorting - data analysis

```
In [2]: import pandas as pd
import seaborn as sns
from matplotlib import pyplot as plt
```

```
In [3]: df=pd.read_excel("President_dataset.xlsx")
```

```
In [9]: df.drop('order',axis=1,inplace=True)
```

```
In [15]: df.describe()
```

```
Out[15]:
```

	height(cm)
count	42.000000
mean	179.738095
std	7.015869
min	163.000000
25%	174.250000
50%	182.000000
75%	183.000000
max	193.000000

```
In [16]: df.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 42 entries, 0 to 41
Data columns (total 2 columns):
name          42 non-null object
height(cm)    42 non-null int64
```

```
dtypes: int64(1), object(1)
memory usage: 752.0+ bytes
```

```
In [18]: df.shape
```

```
Out[18]: (42, 2)
```

```
In [11]: df.sort_values('height(cm)',ascending=True).head()
```

```
Out[11]:
```

	name	height(cm)
3	James Madison	163
7	Martin Van Buren	168
21	Benjamin Harrison	168
1	John Adams	170
22	William McKinley	170

Shortests President are James Madison, Martin Van Buren, Benjamin Harrison

```
In [13]: df.sort_values('height(cm)',ascending=False).head()
```

```
Out[13]:
```

	name	height(cm)
15	Abraham Lincoln	193
33	Lyndon B. Johnson	193
0	George Washington	189
2	Thomas Jefferson	189
39	Bill Clinton	188

Tallest President are Abraham Lincoln, Lyndon B Johnson, George Washington