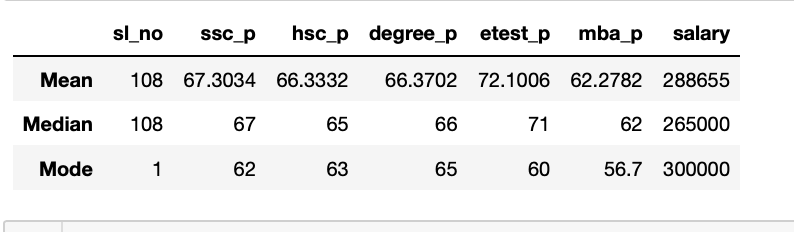
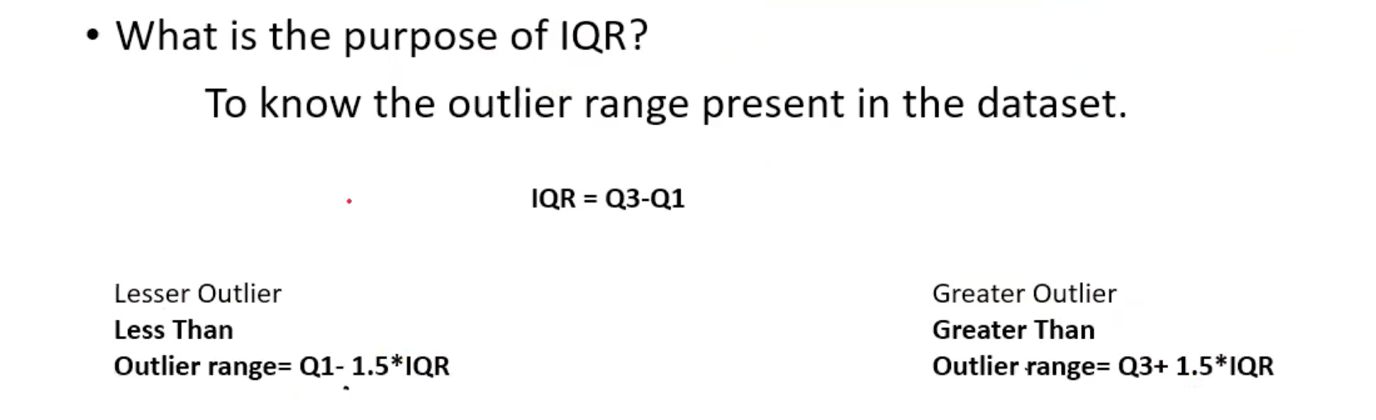
1. **Mean, Median & Mode (MMM) Table Inference**

****

**Inference Tabulation:**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Index** | **Mean** | **Median** | **Outlier Status** | **Mode** |
| **Ssc\_p** | The average ssc pass mark of 215 students = 67.3034 | The average ssc pass mark of 215 students excluding outliers = 67 | Not present | Most repeated pass mark secured in ssc among 215 students= 62 |
| **Take away** | **Average performance in Ssc** | | - | - |
| **Hsc\_p** | The average hsc pass mark of 215 students = 66.3332 | The average hsc pass mark of 215 students excluding outliers = 65 | Present  (Lesser & Greater) | Most repeated pass mark secured in hsc among 215 students= 63 |
| **Take away** | **Average performance in Hsc** | | - | - |
| **Degree\_p** | The average degree pass mark of 215 students= 66.3702 | The average degree pass mark of 215 students excluding outliers = 66 | Present  (Greater) | Most repeated pass mark secured in degree among 215 students= 65 |
| **Take away** | **Average performance in degree graduation** | | - | - |
| **Etest\_p** | The average e-test pass mark of 215 students = 72.1006 | The average e-test pass mark of 215 students excluding outliers = 71 | Not Present | Most repeated pass mark secured in e-test among 215 students= 60 |
| **Take away** | **Above average performance in Etest** | | - | - |
| **Mba\_p** | The average MBA pass mark of 215 students = 62.2782 | The average MBA pass mark of 215 students excluding outliers = 62 | Not present | Most repeated pass mark secured in MBA among 215 students= 56.7 |
| **Take away** | **Average performance in MBA** | | - | - |
| **Salary** | The average salary of 215 students = RS.2,88,655 | The average salary of 215 students excluding outliers= RS.2,65,000 | Present  (Greater) | Most repeated salary package received among 215 students= RS.3,00,000 |
| **Take away** | **Among 215 students, most of the students receive RS.3,00,000 as their salary.** | | | |

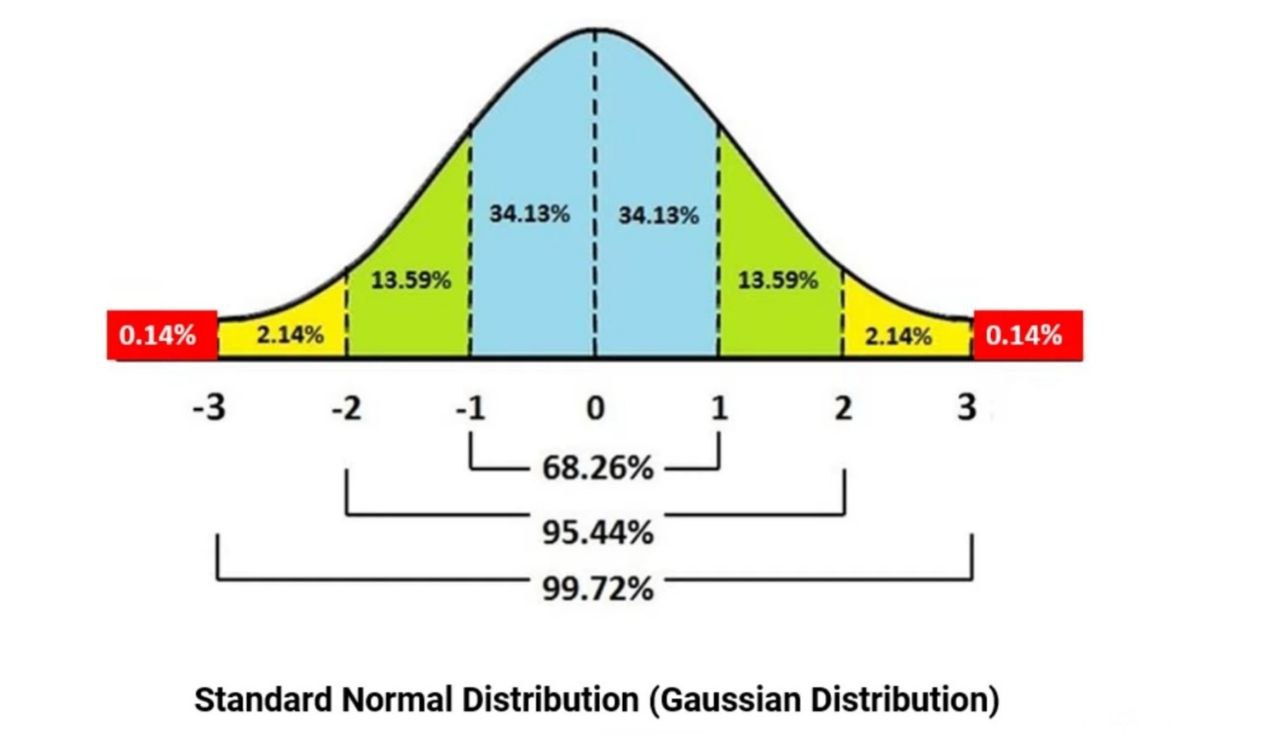
1. **Formula Reasoning:**

****

**Question :** Why 1.5 scale is multiplied with IQR for outliers?

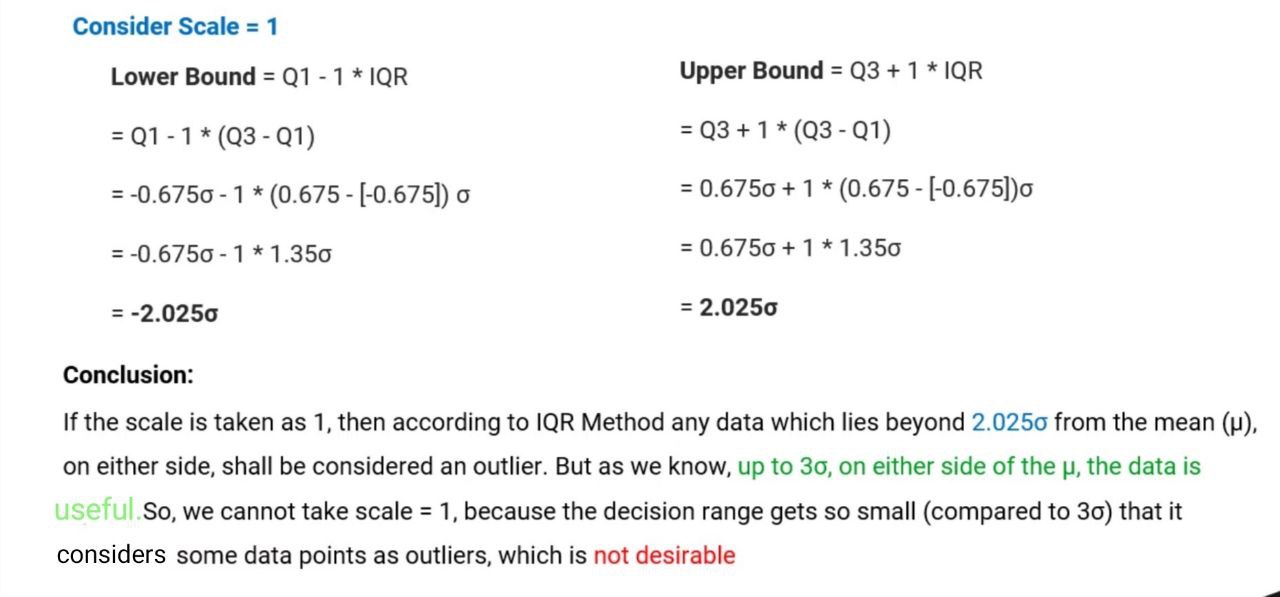
**Reason:**

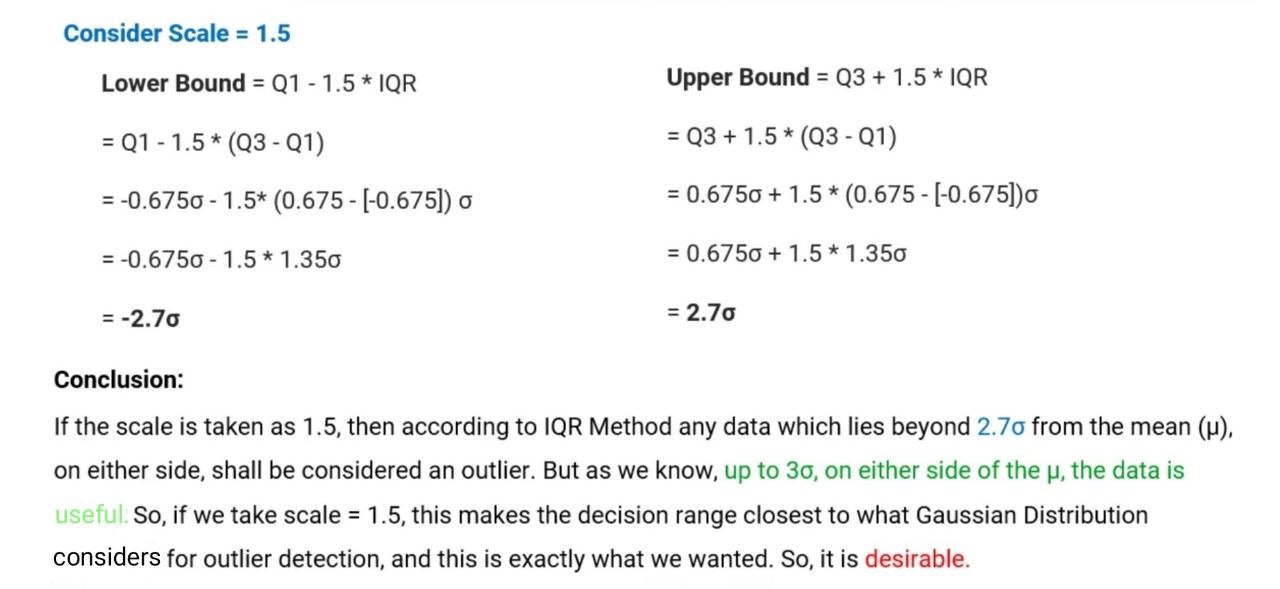
Consider the Gaussian distribution decision range for outlier detection as shown below.

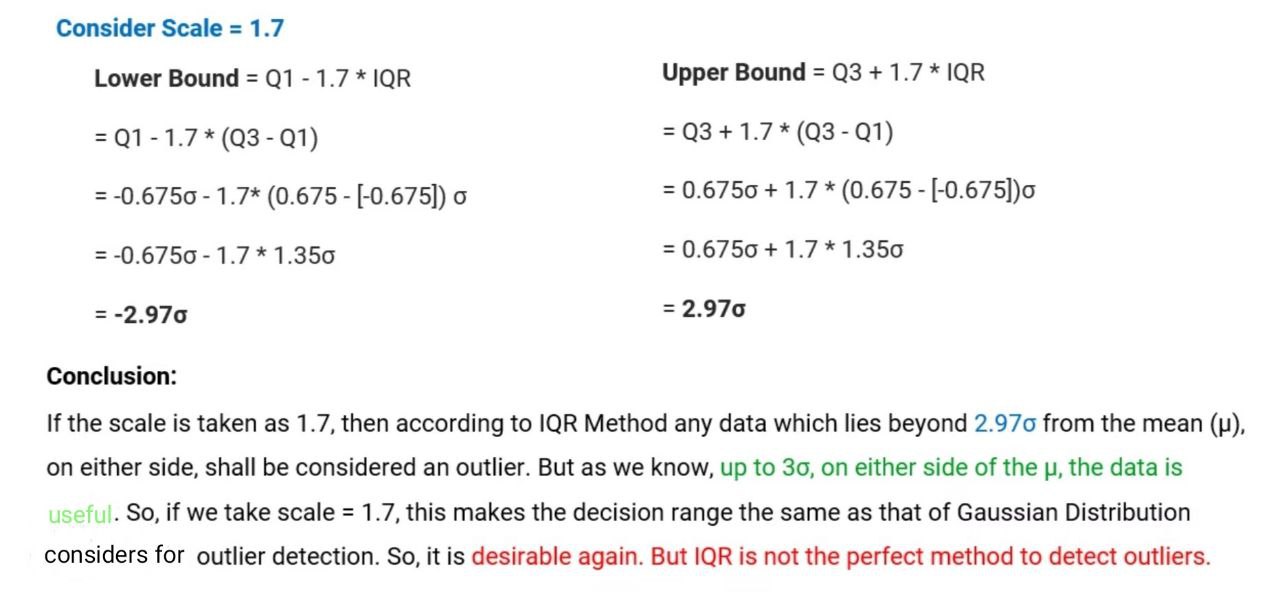


Here we are looking for the data which is useful for us which lies within the range mean plus or minus 3 sigma data points. Those data points which lies beyond this range is considered as outlier.

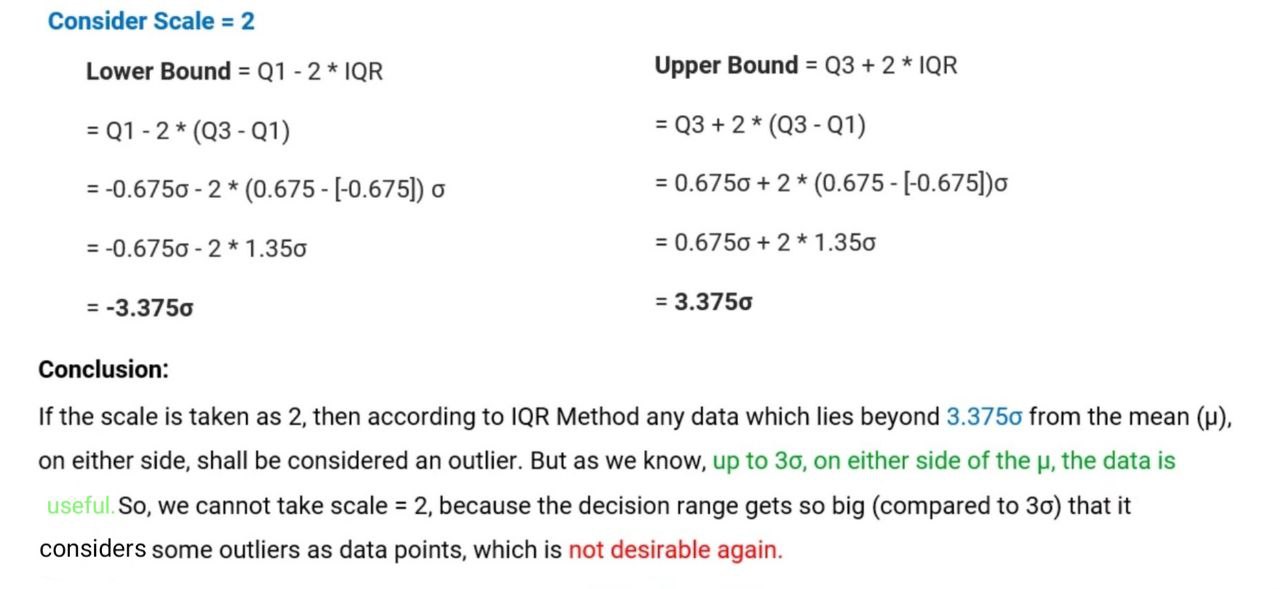
To achieve this let’s try to calculate the approximate scale value which has to be multiplied with IQR so that we end up with ±3σ range.





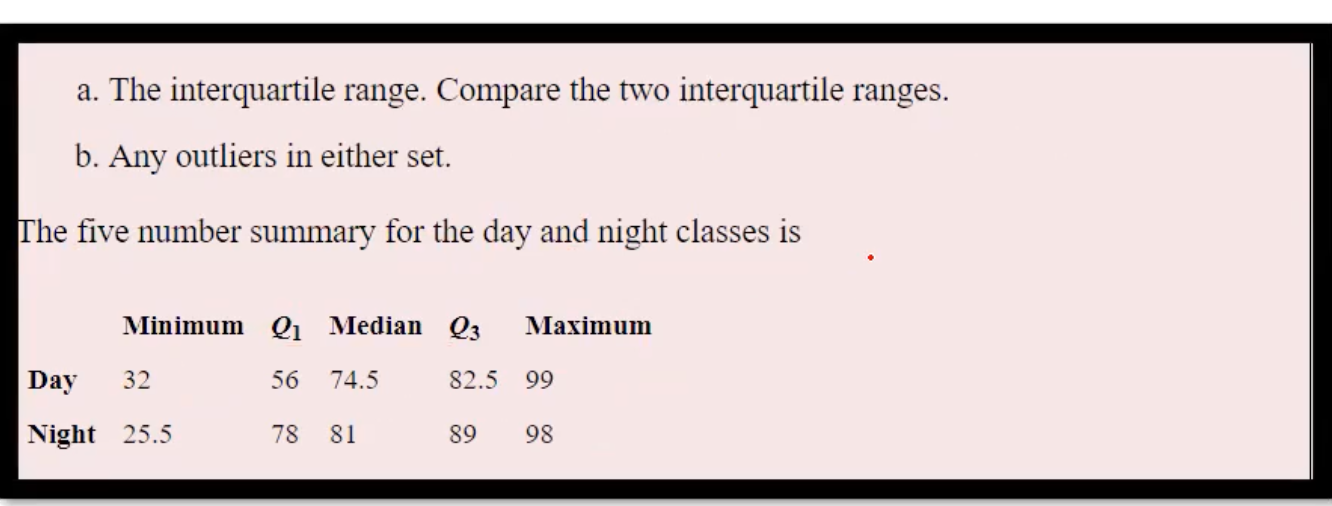


To keep the tolerance we choose 1.5 scale on safer side though 1.7 scale suits the best.



1. **Problem on IQR**

**Question :**

****

**Solution :**

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **Classes** | **Q1** | **Q3** | **IQR= Q3- Q1** | **Lesser outlier range =**  **Q1-1.5\*IQR** | **Greater outlier range =**  **Q3 + 1.5\*IQR** | **Upper & Lower Range** |
| **Day** | 56 | 82.5 | 26.5 | 16.25 | 122.25 | **Range:(16.25, 122.25)**  No values in the row named day is outside the range. So no outliers are present. |
| **Night** | 78 | 89 | 11 | 61.5 | 105.5 | **Range:(61.5, 105.5)**  Value 25.5 in the row named night is outside the range. So outlier is present.  **Type of outlier :**  **Low range outlier** |

**4.Kurtosis & Skewness Table Inference**

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
|  | **sl\_no** | **ssc\_p** | **hsc\_p** | **degree\_p** | **etest\_p** | **mba\_p** | **salary** |



**Inference:**

**Kurtosis: (Tells about the Tailness)**

Kurtosis value for all the columns starting from ssc\_p to salary is <3 which means the type of kurtosis is platykurtic( Negative Excess Kurtosis).

Which in turn gives the information that the number of students falling under below average & above average (to some extent) whereas number of students falling under average (to better extent) in terms of exam performance & receiving salary, where the most population of students fall under average criteria.

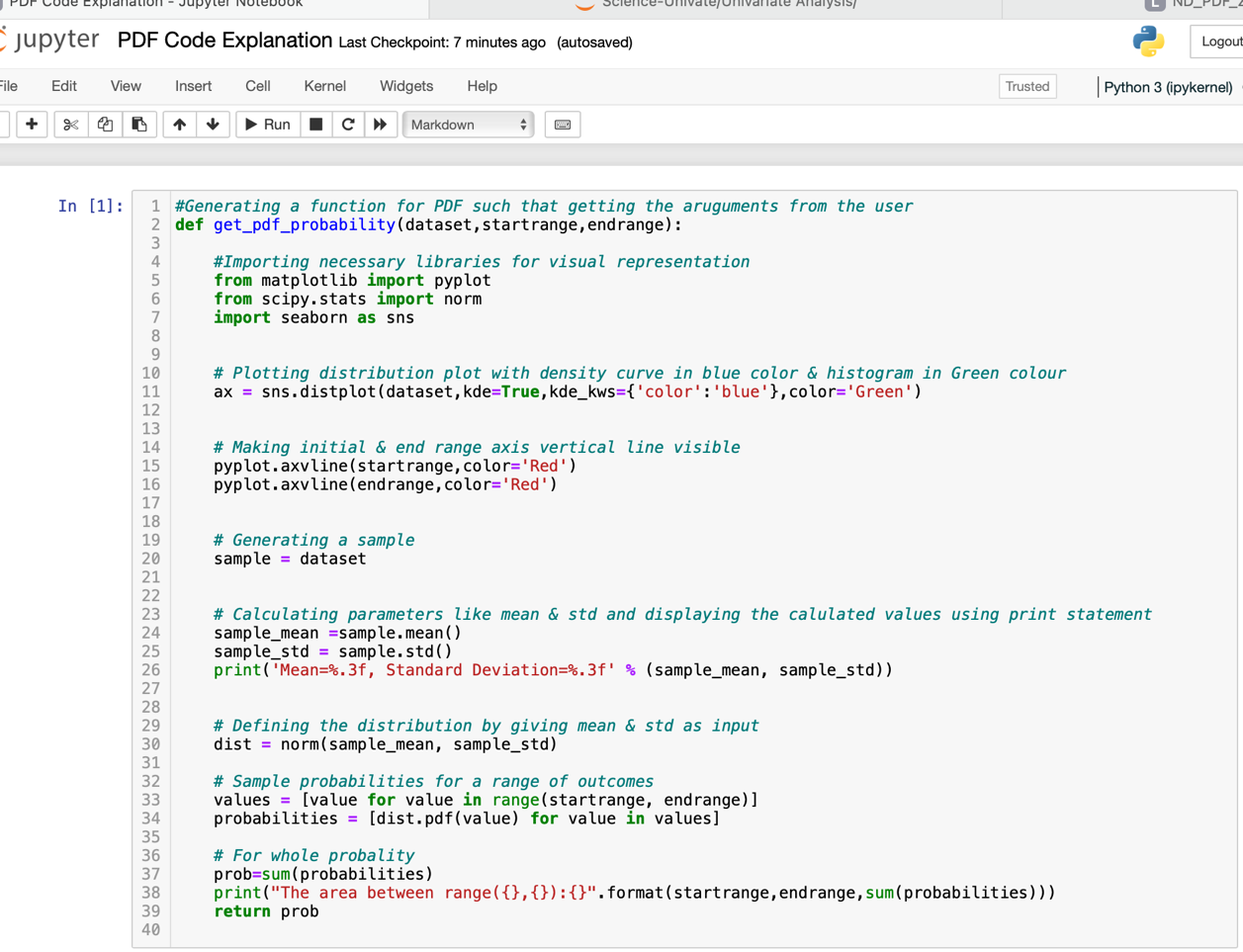
**Skewness: (Tells us about Asymmetry from the normal distribution)**

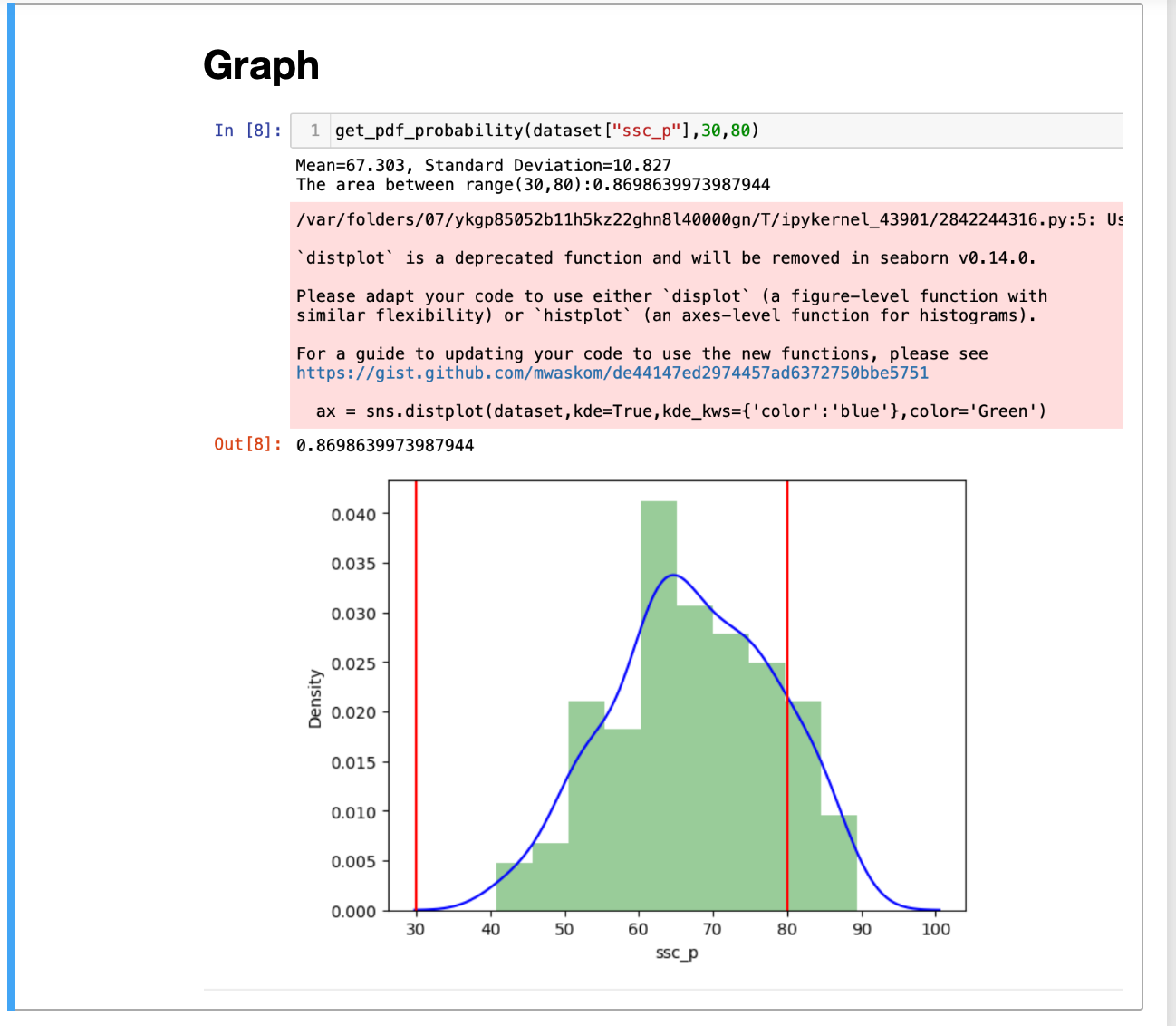
Ssc\_p alone exhibits negative skewness.i.e, Above average scorers are more in count when compared to below average scorers.

Hsc\_p, Degree\_p, Etest\_p & MBA \_p variables exhibits positive skewness. i.e Above average scorers are lesser in count when compared to below average scorers.

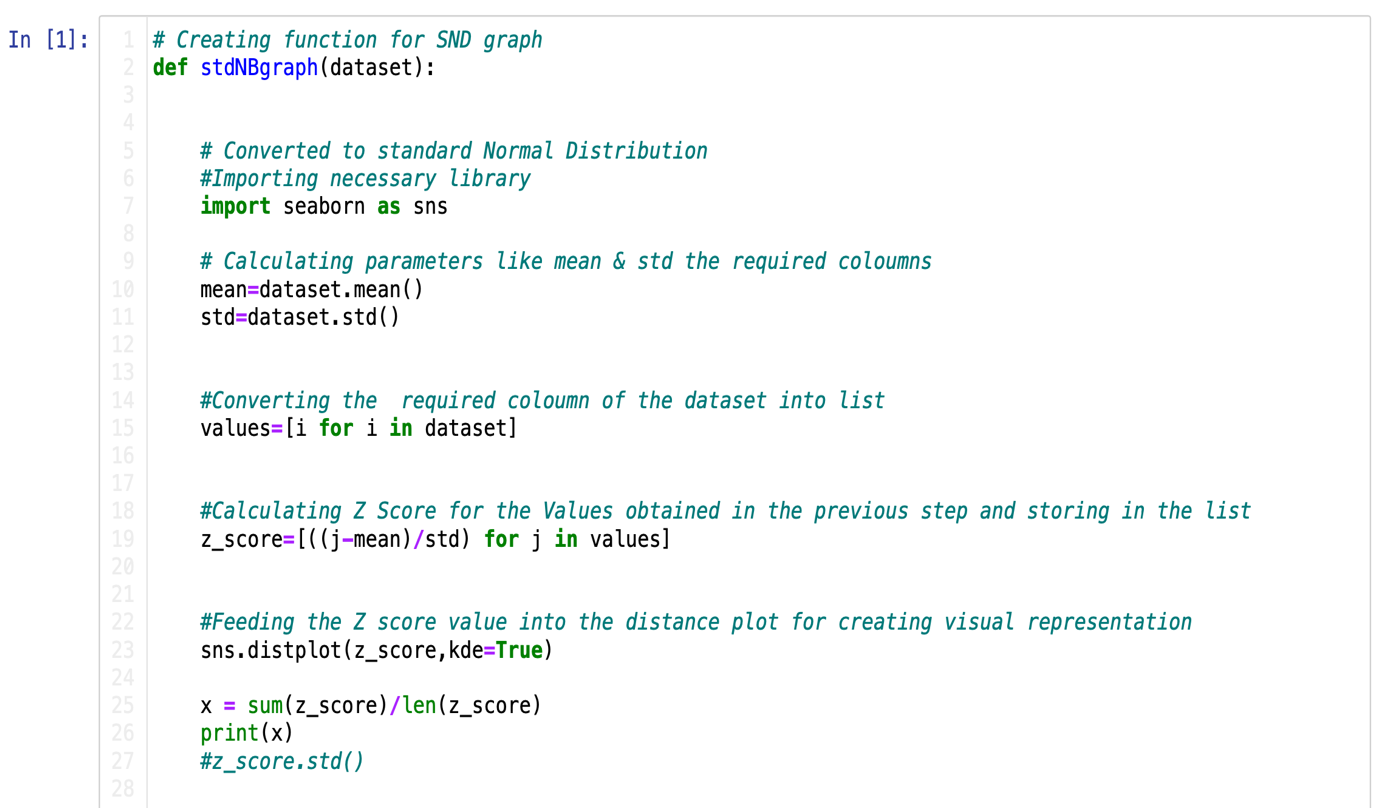
Salary variable also exhibits positive skewness. i.e Above average salary receivers are lesser in count when compared to below average salary receivers.

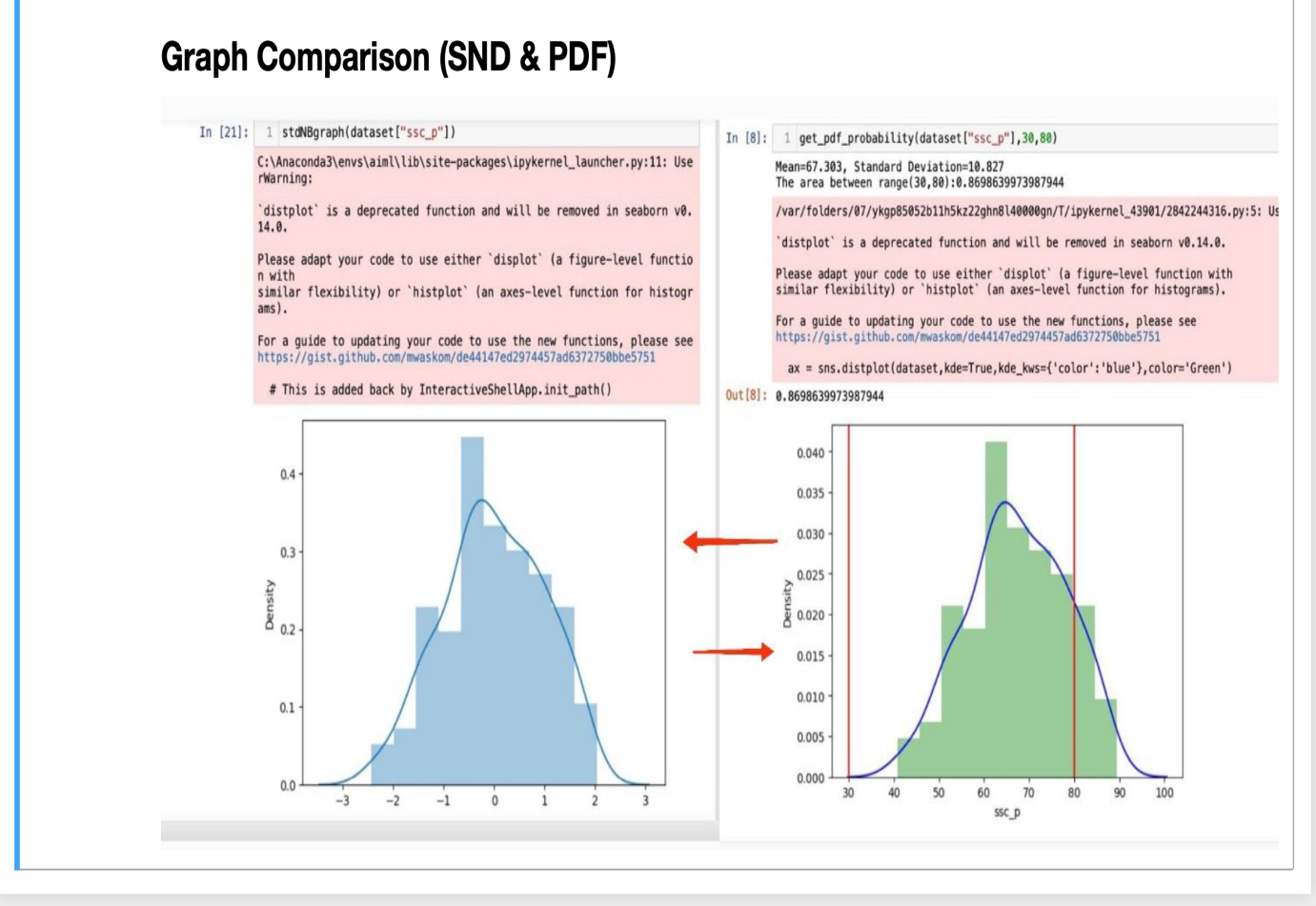
**5. Probability Distribution Function (PDF) Code Explanation:**





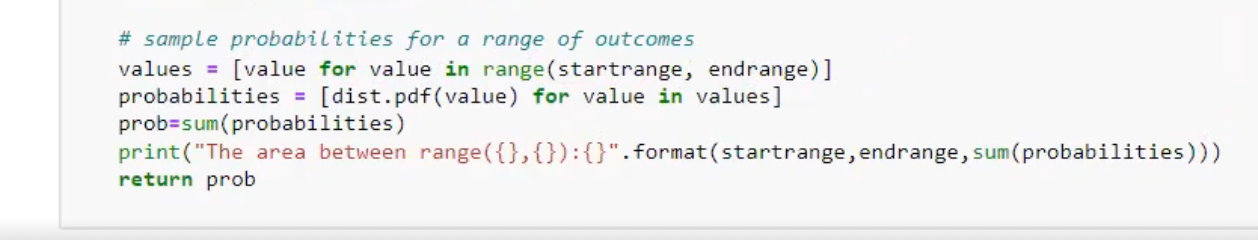
**6. Standard Normal Distribution (SND) Code Explanation:**

****

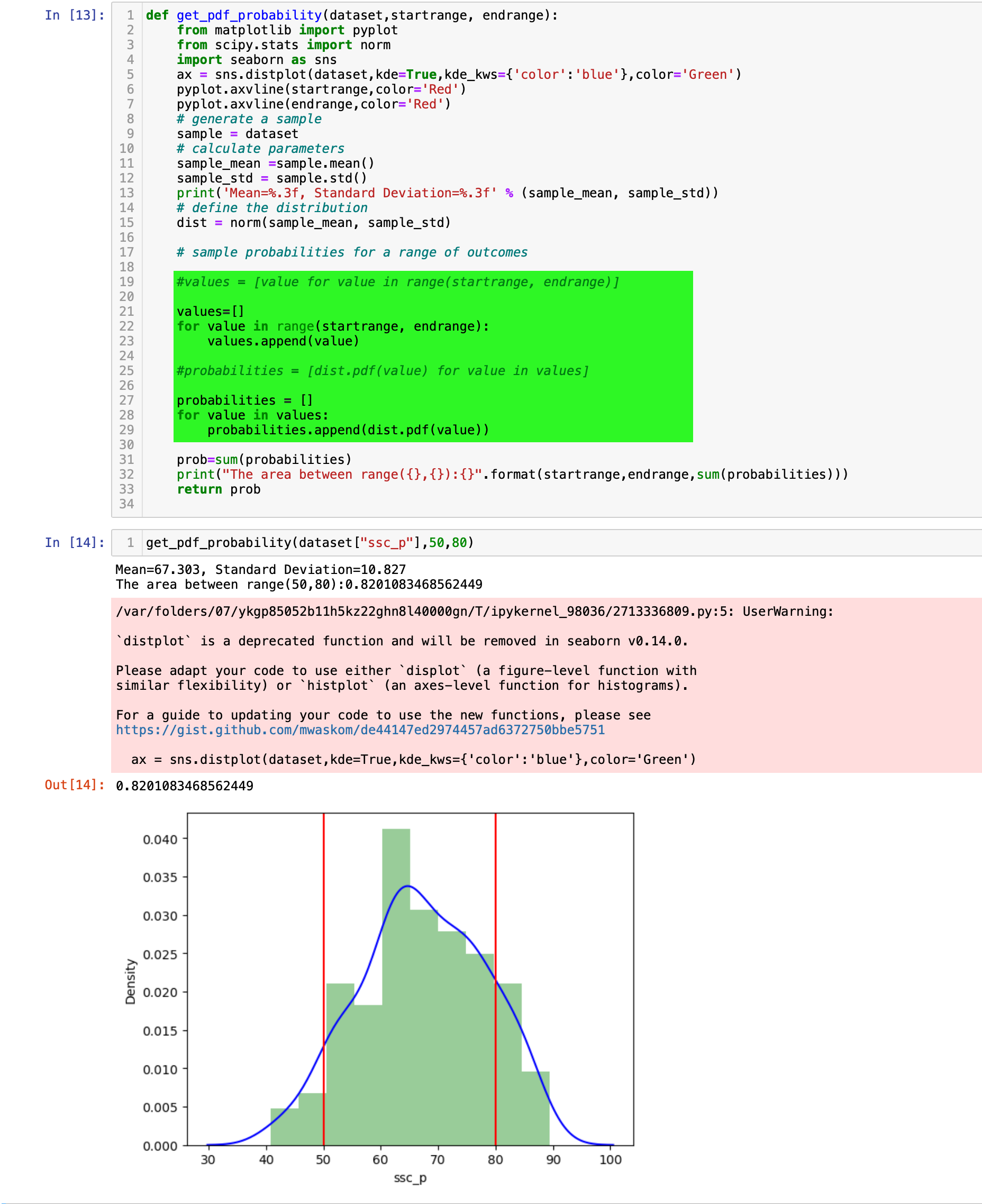
****

**7. For Loop in Traditional Way**

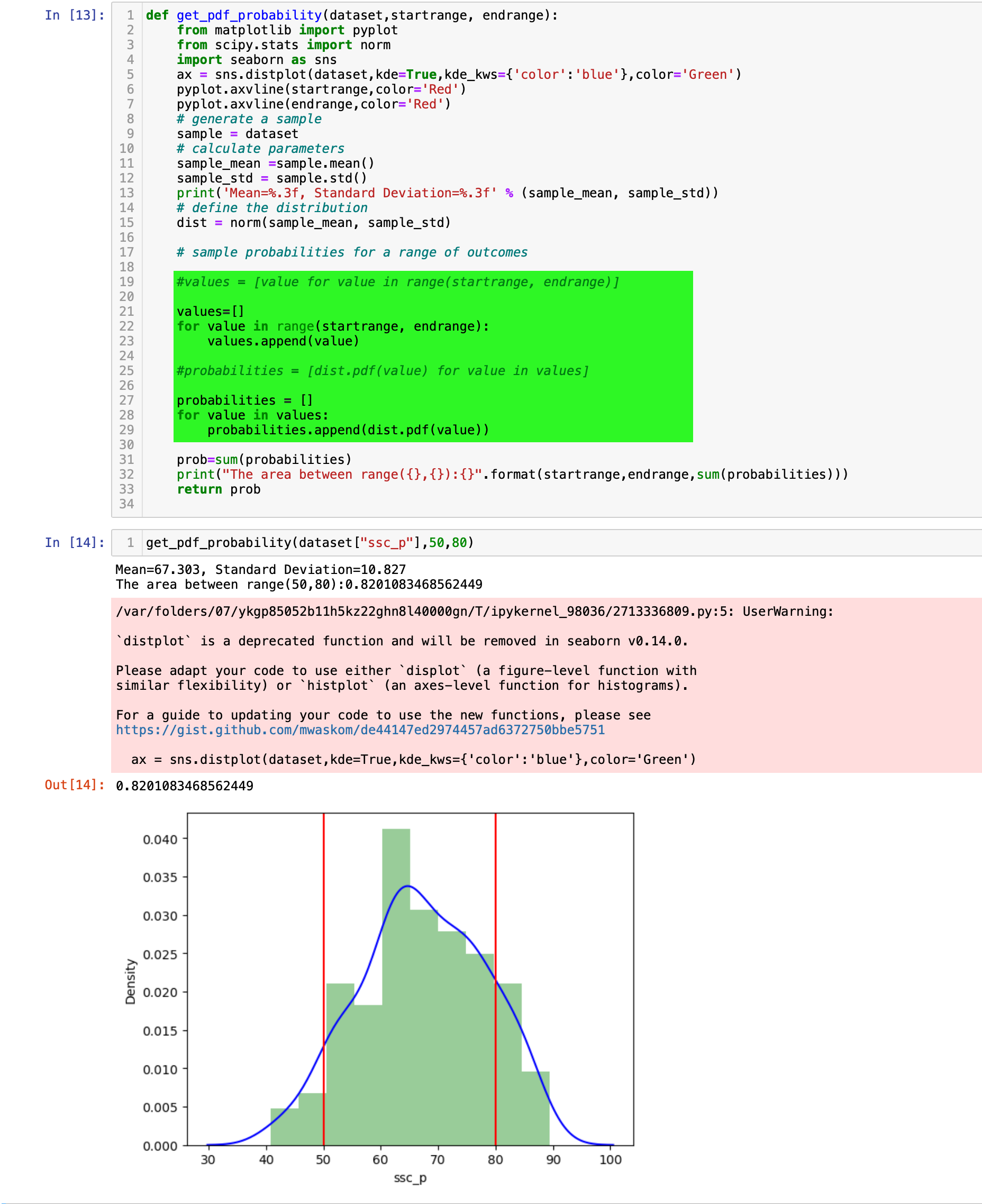
**One Liner For Loop:**

****

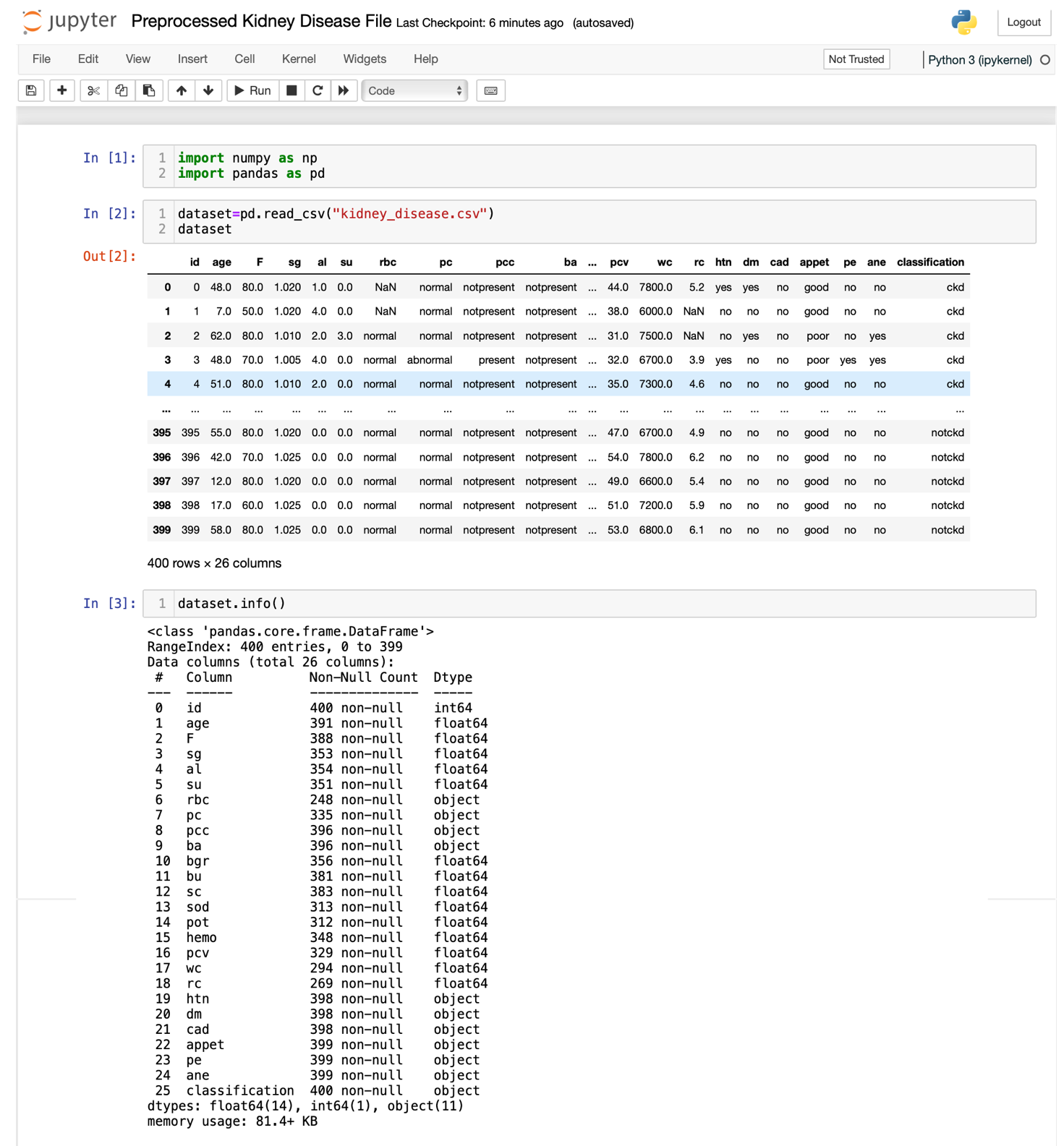
**Tradition Way For Loop:**

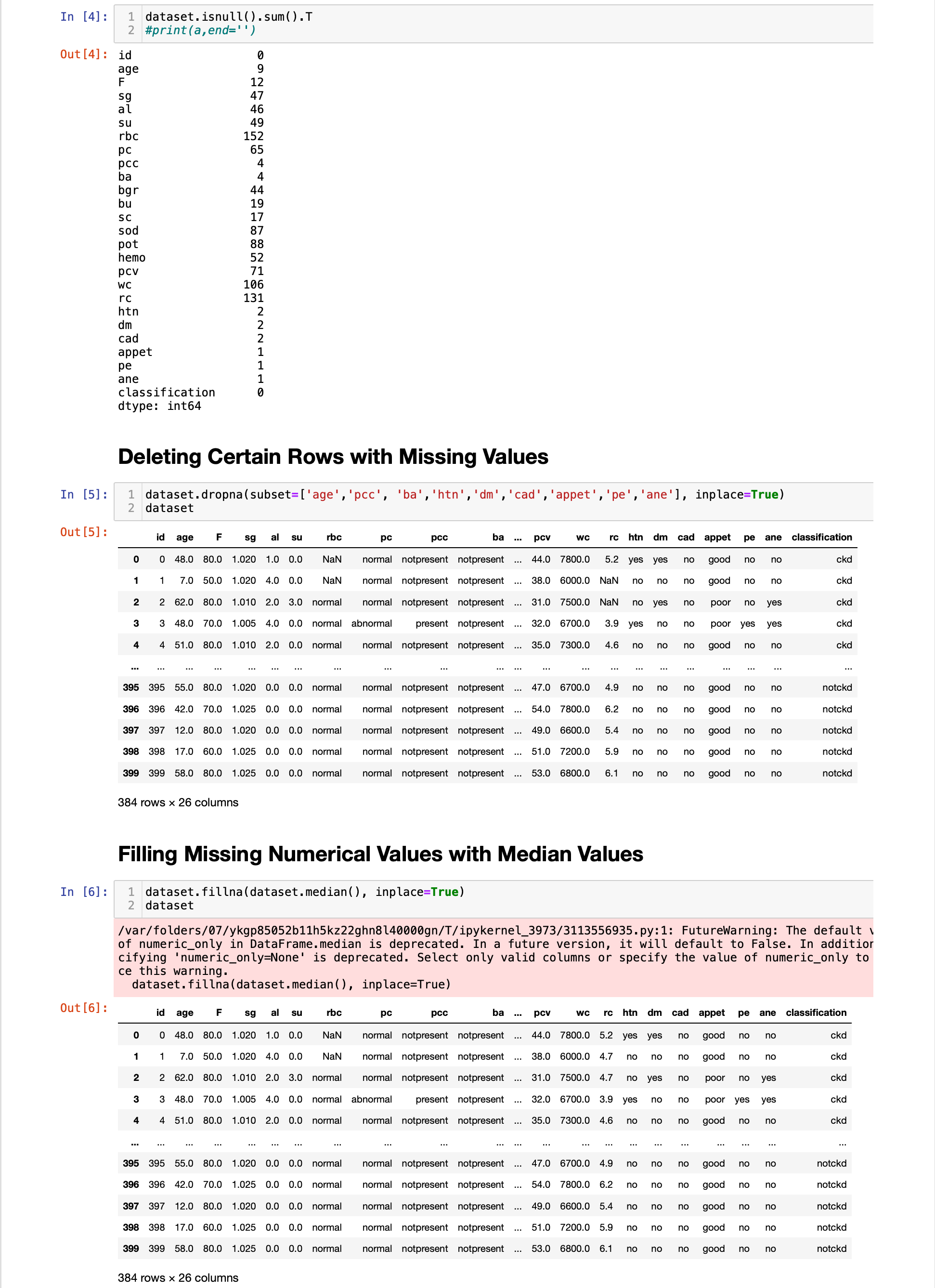
****

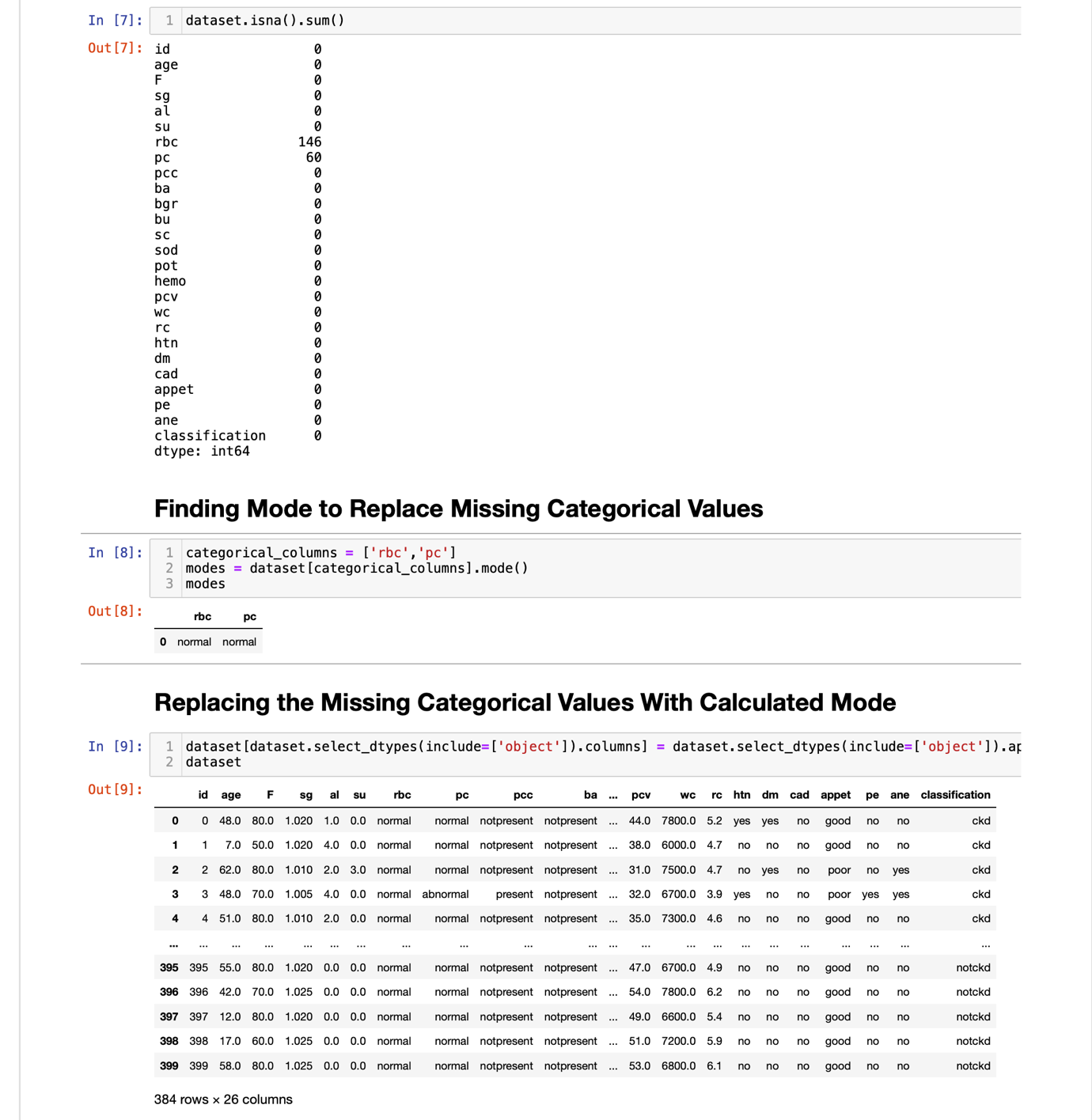
**Proof:**

****

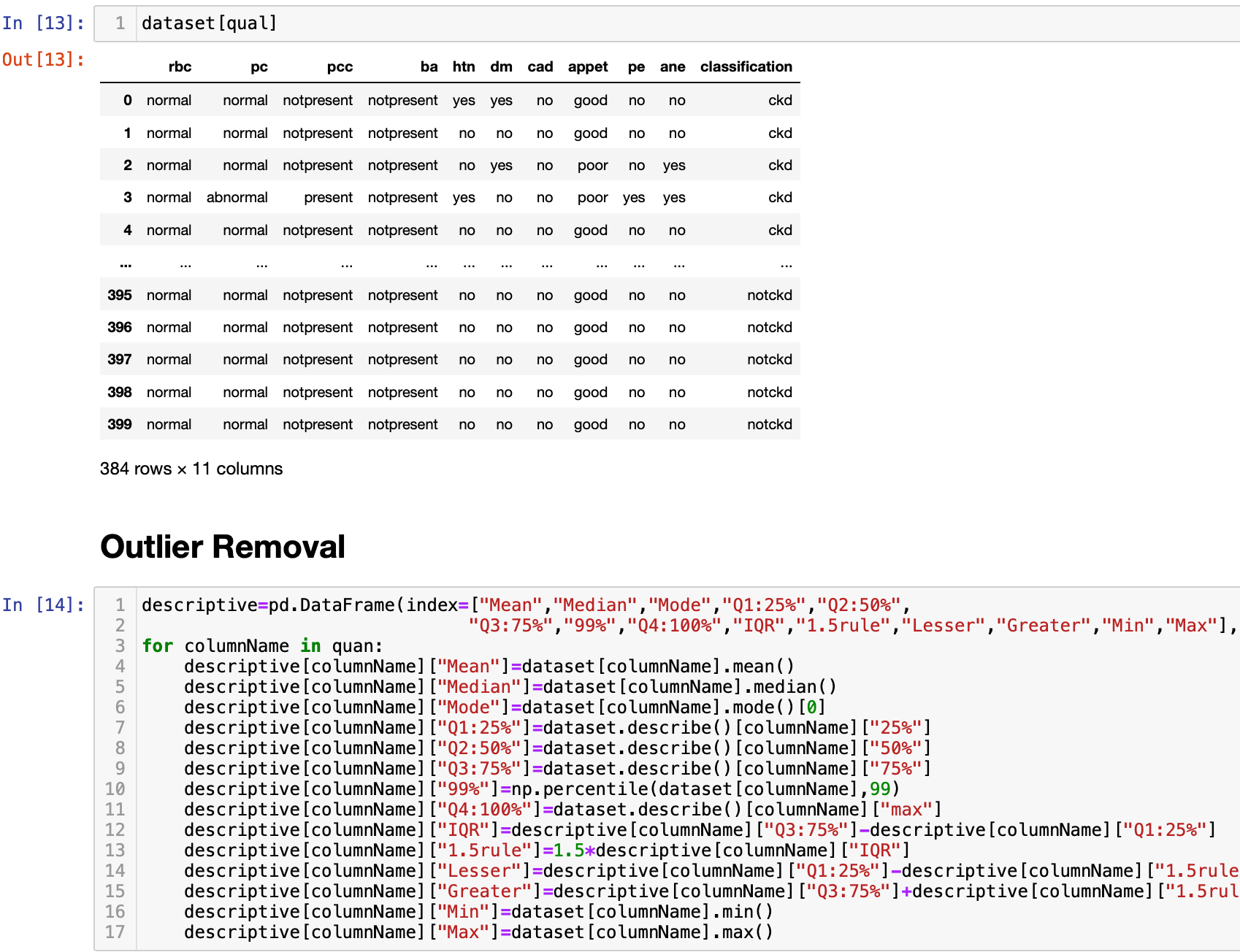
**8.CKD File Preprocessing**

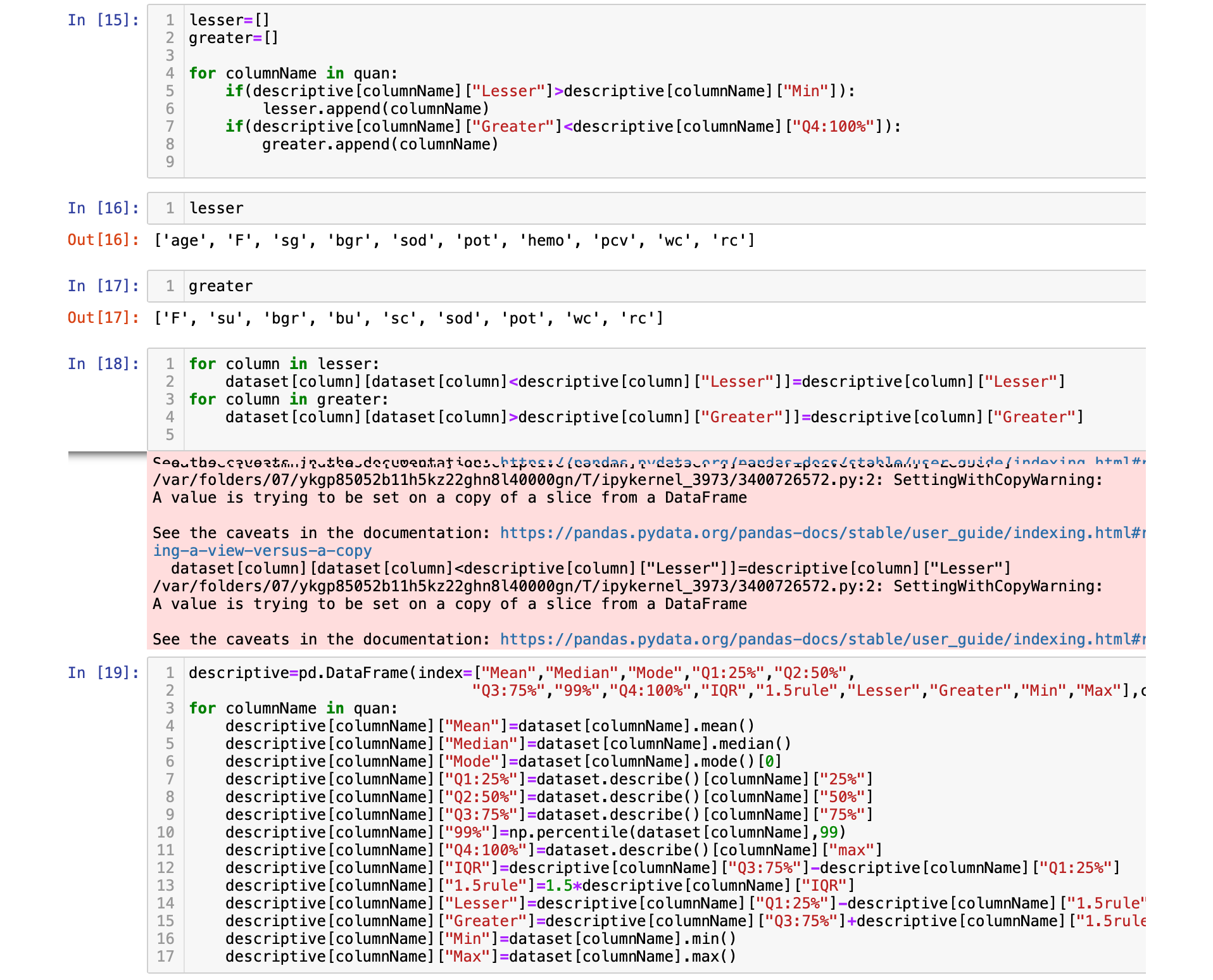
****

****

****

****

****

****

****

**---- The End----**