**J.Vineetha**

**G34 - Python&ML**

1.Open the [Teaching Ratings data](https://courses.cognitiveclass.ai/assets/courseware/v1/8b5e886922d6d1b5cf057fd0ff3b974a/asset-v1:IBM+ST0101EN+v1+type@asset+block/teachingratings.sav) in SPSS and switch to variable view. Consider the following variables:  **age, gender, beauty, eval, tenure, students**. Thought questions: Can you identify which variables are **continuous variables** and which ones are **categorical variables**? Are any of the variables in the above list of **ordinal** type?

**Continuous variables** (Scale/Interval/Ratio):

1. age
2. beauty
3. eval
4. students

**Categorical variables** (Nominal/Ordinal):

* 1. gender – Nominal (e.g., male/female)
  2. tenure – Nominal (tenured or not)

**Ordinal variable** (ranked categories):

In this dataset, **none of the six listed** are ordinal. But if tenure had levels like “assistant, associate, full professor”, that might be ordinal.

2.Can you identify whether the Teaching Ratings data is a time series, cross-sectional, and/or multivariate data set?

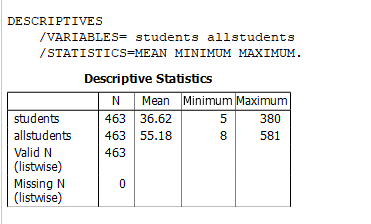
Cross-sectional and Multivariate

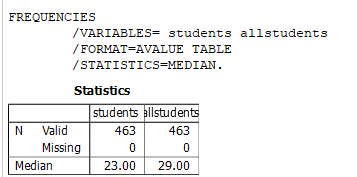
3.Does the Teaching Ratings data set represents information on an entire population or just a sample?

Usually, data like this is a **sample** taken from a larger group (e.g., a group of professors from one university).

So, this dataset **represents a sample**, **not an entire population**.

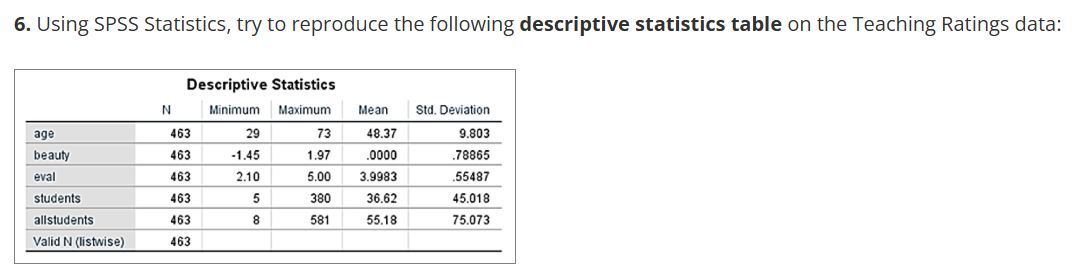
**4.**  Using SPSS Statistics, find the **mean**, **median**, **minimum**, and **maximum** values for **students** and **allstudents** in the Teaching Ratings data.

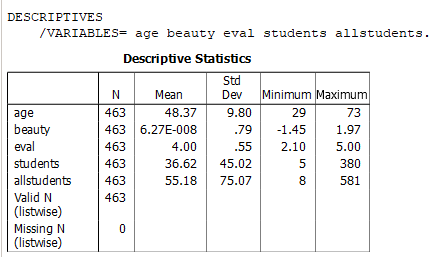




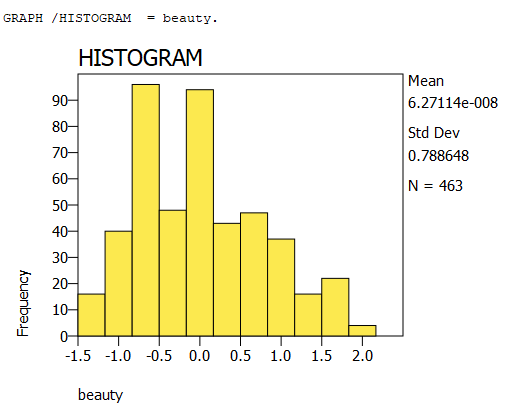
**5.** What is the mathematical relationship between **variance** and **standard deviation**?







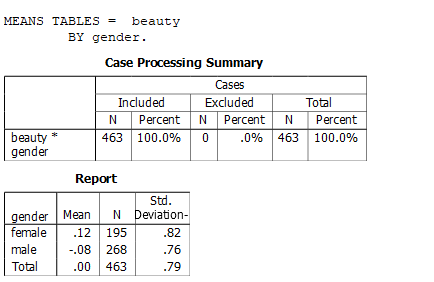
**7.** Create a histogram of **beauty** and briefly comment on the distribution of data (as a thought question).



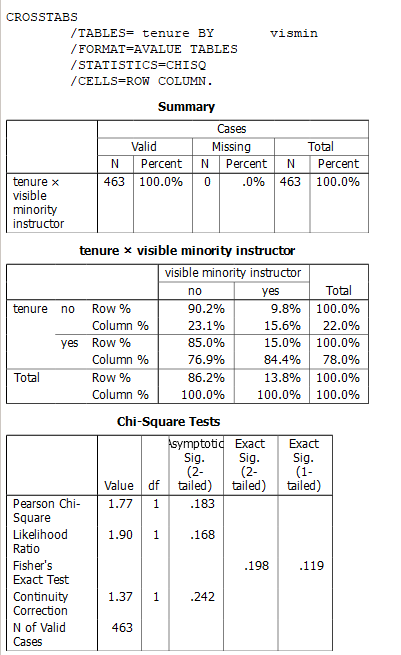
COMMENT:

The beauty scores look a bit uneven. Most of the values are around 0, but there are more people with lower scores than higher ones. A few people have very high beauty scores, which makes the graph stretch more to the right side. So, the shape of the graph is slightly skewed to the right. It’s not perfectly even on both sides.

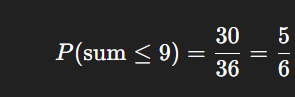
**8.** Does average **beauty** score differ by **gender** in the Teaching Ratings data? Produce averages and standard deviations for each of the two variables.



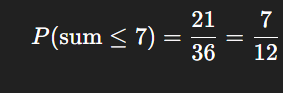
**9.** Does **tenure** status differ by **vismin** (visible minority) status? Produce cross tabulations explaining what percentage of visible minorities are tenured.



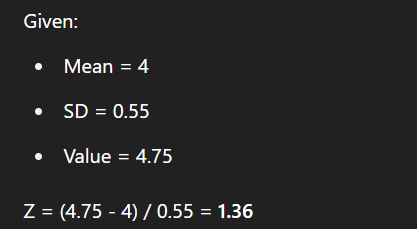
**10.** What is the probability of rolling two dice and getting 9 or less?



**11.** What is the probability of rolling two dice and getting 7 or less?



**12.** With an average teaching evaluation score of 4 and standard deviation of 0.55, what is the probability of getting a teaching evaluation of greater than 4.75?



Using Z-Table, for Z=1.36

We know that Z-Table have always left region values.

So to get right region we need to subtract left region from 1.

