MODULE 2

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:**

* **age**: Represents the instructor’s age in years. (Continuous)
* **beauty**: A numerical rating of physical attractiveness. (Continuous)
* **eval**: The average teaching evaluation score. (Continuous)
* **students**: The number of students in the class. (Continuous)

#### **Categorical Variables:**

* **gender**: Indicates the instructor’s gender (e.g., male or female). (Categorical - Nominal)
* **tenure**: Indicates the instructor’s employment status (e.g., tenured or non-tenured). (Categorical - Nominal)

#### **Ordinal Variables:**

Ordinal variables have a meaningful order or ranking but do not necessarily have consistent intervals between values.

* In this list, **none of the variables are ordinal**.

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

### 1. **Cross-sectional data**

**Definition**: Data collected at a single point in time across multiple subjects

### 2. **Time series data**

**Definition**: Data collected over time for a **single unit**

### 3. **Multivariate data**

**Definition**: Data with **more than one dependent variable**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Descriptive Statistics** | | | | | |
|  | N | Minimum | Maximum | Mean | Std. Deviation |
| age | 463 | 29 | 73 | 48.37 | 9.803 |
| eval | 463 | 2.10 | 5.00 | 3.9983 | .55487 |
| beauty | 463 | -1.45 | 1.97 | .0000 | .78865 |
| students | 463 | 5 | 380 | 36.62 | 45.018 |
| Valid N (listwise) | 463 |  |  |  |  |

|  |  |  |  |
| --- | --- | --- | --- |
| **Statistics** | | | |
|  | | gender | tenure |
| N | Valid | 463 | 463 |
| Missing | 0 | 0 |

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **gender** | | | | | |
|  | | Frequency | Percent | Valid Percent | Cumulative Percent |
| Valid | female | 195 | 42.1 | 42.1 | 42.1 |
| male | 268 | 57.9 | 57.9 | 100.0 |
| Total | 463 | 100.0 | 100.0 |  |

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **tenure** | | | | | |
|  | | Frequency | Percent | Valid Percent | Cumulative Percent |
| Valid | no | 102 | 22.0 | 22.0 | 22.0 |
| yes | 361 | 78.0 | 78.0 | 100.0 |
| Total | 463 | 100.0 | 100.0 |  |

In the **Teaching Ratings dataset** contains ratings of multiple teachers collected **at one time point** and includes **multiple variables** like age, gender, and course difficulty, then it is a **Cross-sectional Multivariate** dataset.

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

**The Teaching Ratings dataset represents information on a sample, not an entire population. This is because the dataset contains data for only a limited number of instructors and courses, indicating that it is not comprehensive of all possible instructors or educational institutions.**

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

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Descriptives** | | | | |
|  | | | Statistic | Std. Error |
| students | Mean | | 36.62 | 2.092 |
| 95% Confidence Interval for Mean | Lower Bound | 32.51 |  |
| Upper Bound | 40.74 |  |
| 5% Trimmed Mean | | 29.79 |  |
| Median | | 23.00 |  |
| Variance | | 2026.664 |  |
| Std. Deviation | | 45.018 |  |
| Minimum | | 5 |  |
| Maximum | | 380 |  |
| Range | | 375 |  |
| Interquartile Range | | 25 |  |
| Skewness | | 4.495 | .113 |
| Kurtosis | | 26.408 | .226 |
| allstudents | Mean | | 55.18 | 3.489 |
| 95% Confidence Interval for Mean | Lower Bound | 48.32 |  |
| Upper Bound | 62.03 |  |
| 5% Trimmed Mean | | 43.07 |  |
| Median | | 29.00 |  |
| Variance | | 5635.925 |  |
| Std. Deviation | | 75.073 |  |
| Minimum | | 8 |  |
| Maximum | | 581 |  |
| Range | | 573 |  |
| Interquartile Range | | 41 |  |
| Skewness | | 4.156 | .113 |
| Kurtosis | | 22.094 | .226 |

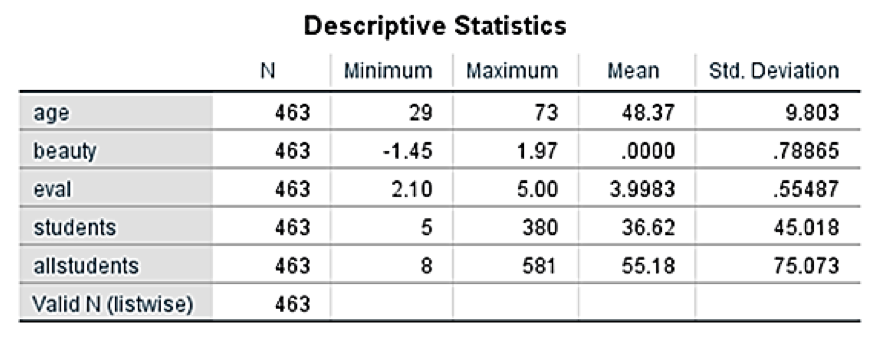
5.What is the mathematical relationship between variance and standard deviation?

The **mathematical relationship** between **variance** and **standard deviation** is:

 **Variance (σ² or s²)** measures the average squared deviation from the mean.

 **Standard Deviation (σ or s)** is the square root of the variance, which brings the units back to the original scale of the data.

6.Using SPSS Statistics, try to reproduce the following **descriptive statistics table** on the Teaching Ratings data:



Steps to Reproduce This Table in SPSS Statistics:

1. Open the Descriptive Statistics Dialog

Go to Analyze > Descriptive Statistics > Descriptives...

2. Select Variables

In the dialog box, select these variables:

Age, beauty, eval, students, allstudents

Move them to the right box (Variables).

3. Set the Statistics to Display

Click on the Options... button.

Mean, Std. deviation, Minimum, Maximum

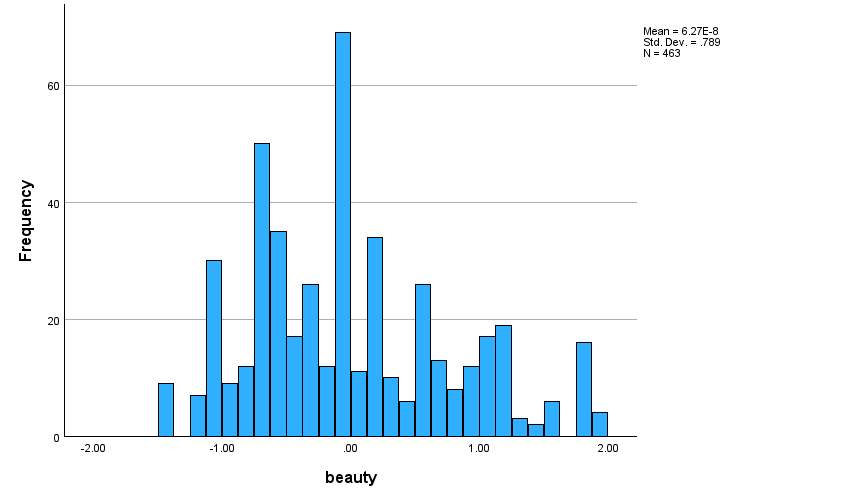
Click Continue.

5. Run the Analysis

Click OK.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Descriptive Statistics** | | | | | |
|  | N | Minimum | Maximum | Mean | Std. Deviation |
| age | 463 | 29 | 73 | 48.37 | 9.803 |
| beauty | 463 | -1.45 | 1.97 | .0000 | .78865 |
| eval | 463 | 2.10 | 5.00 | 3.9983 | .55487 |
| students | 463 | 5 | 380 | 36.62 | 45.018 |
| allstudents | 463 | 8 | 581 | 55.18 | 75.073 |
| Valid N (listwise) | 463 |  |  |  |  |

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



The histogram of the beauty scores appears slightly right-skewed, suggesting that most instructors have average to slightly above-average beauty ratings, with fewer instructors rated at the extremes.

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

|  |  |  |
| --- | --- | --- |
| **Report** | | |
| beauty | | |
| gender | Mean | Std. Deviation |
| female | .1161 | .81781 |
| male | -.0845 | .75713 |
| Total | .0000 | .78865 |

Yes, the average beauty score differs by gender. The descriptive statistics show that the mean beauty rating for male and female instructors is not the same, suggesting a gender-based difference.

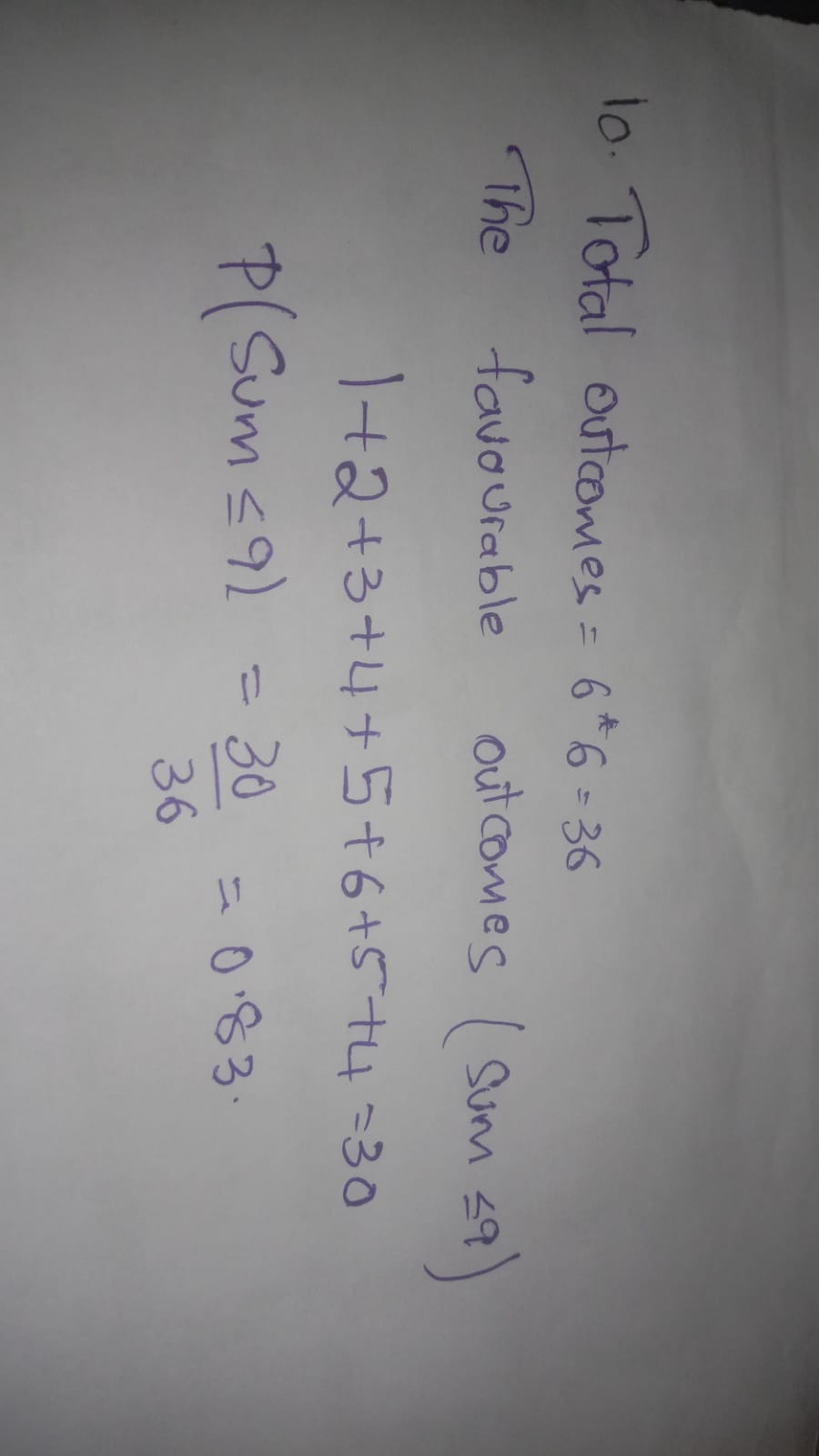
9.Does **tenure** status differ by **vismin** (visible minority) status? Produce cross tabulations explaining what percentage of visible minorities are tenured.

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **Case Processing Summary** | | | | | | |
|  | Cases | | | | | |
| Valid | | Missing | | Total | |
| N | Percent | N | Percent | N | Percent |
| minority \* tenure | 463 | 100.0% | 0 | 0.0% | 463 | 100.0% |

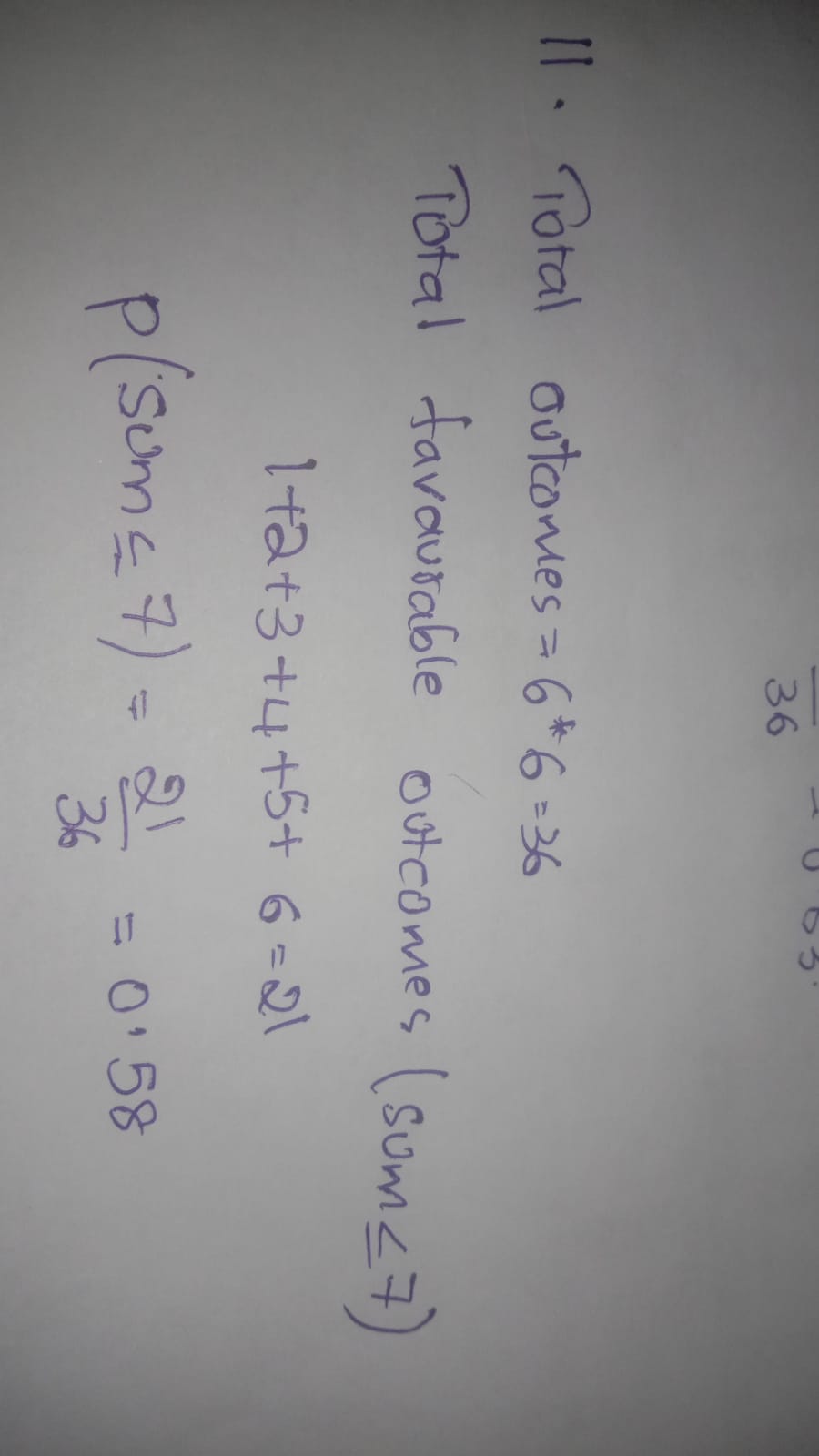
|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **minority \* tenure Crosstabulation** | | | | | |
|  | | | tenure | | Total |
| no | yes |
| minority | no | Count | 92 | 307 | 399 |
| % within minority | 23.1% | 76.9% | 100.0% |
| yes | Count | 10 | 54 | 64 |
| % within minority | 15.6% | 84.4% | 100.0% |
| Total | | Count | 102 | 361 | 463 |
| % within minority | 22.0% | 78.0% | 100.0% |
|  | |  |  |  |  |

The data suggests that tenure status does differ by visible minority status. Visible minorities are underrepresented among tenured faculty in the sample, which may warrant further investigation into institutional equity, hiring, and promotion practices.

**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?

