|  |
| --- |
| 1. **Which data storage format is efficient for handling large volumes of structured data?** *a) .txt b) .csv c) .parquet d) .doc* |
| 1. **Which method is used to reshape data in Pandas from long to wide format?** *a) melt() b) join() c) pivot() d) reshape()* |
| 1. What is the primary purpose of data wrangling? *a) Collecting raw data b) Cleaning and transforming data into a usable format c) Visualizing data d) Storing data in databases* 2. **Which method would best help in summarizing and grouping data?** *a) sort() b) groupby() c) melt() d) fillna()* 3. **Which of the following helps detect outliers in a dataset?** *a) Mean b) Median c) Standard Deviation d) Boxplot* |
| 1. Given a DataFrame *df*with missing values, which function fills missing values with the mean?   *a) df.dropna() b) df.fillna(df.mean())*  *c) df.replace(np.nan, 0) d) df.interpolate()*  7. Which method removes duplicate rows in a DataFrame?  *a) df.drop\_duplicates() b) df.unique()  c) df.remove\_duplicates() d) df.deduplicate()* |
| 1. Noise and anomalies in data can be detected using: a) *Random sampling b) Statistical techniques and visualizations c) Data duplication d) Merging columns* |
|  |
|  |
|  |
|  |
| 8. What does "hello".replace("l", "z") return? *a) "hezzo" b) "hezzzo" c) "hello" d) "heo"* |
| 9. Which code snippet correctly splits a string into words?  *s = "Data Science"*  *a) s.split(" ") b) s.split(",") c) s.split("") d) s.slice(" ")* |
| 10.Which of the following code snippets creates a scatter plot of x vs y with red points?  *a) plt.scatter(x, y, color='red') b) plt.scatter(x, y, c='red') c) plt.scatter(x, y, point\_color='red') d) Both a and b* |
| 11. What does a boxplot represent? *a) Trend over time b) Distribution of numerical data (median, quartiles, outliers) c) Correlation between variables d) Categorical counts* |
| 12. How can you add a title to a Matplotlib plot? *a) plt.title("My Plot") b) plt.label("My Plot") c) plt.heading("My Plot") d) plt.name("My Plot")* |
| 13. The interquartile range (IQR) is the distance between:  *a) Min and Max b) Q1 and Q3*  *c) Median and Q3 d) Max and Q1* |
| 14. Which Seaborn function creates a scatter plot with a regression line?  *a) sns.scatterplot() b) sns.regplot()*  *c) sns.lineplot() d) sns.pairplot()*  *15****.* Which of the following is used to control the ticks on an axis?** *a) plt.tick\_params() b) plt.label() c) plt.grid() d) plt.annotate()* |
| 16. **To save a plot into a file like PNG or PDF, which command is used?** *a) saveplot() b) plt.save() c) plt.savefig() d) plt.file()*  17*.* **Which Seaborn function is used to create a scatter plot?** *a) scatter() b) sns.scatterplot() c) sns.lineplot() d) sns.boxplot()*  18***.* A Histogram is best used to:** *a) Analyze correlation b) Visualize frequency distribution c) Compare categories d) Display 3D data* |
| 19. **Which of the following can NOT be changed when styling your plots?** *a) Line width b) Background color c) X-Y data points d) Font size*  20.**How do you create multiple plots in one figure in Matplotlib?** *a) Using subplot() b) Using groupby() c) Using subplots() d) Using plotgrid()* |
|  |