Introduction:

This document describes the process and rules for writing questions for given methods.
You need to go through the examples and rules in this document before writing questions.

Rules:

1. **The questions must be clear and understandable.**

2. **Include the table and columns in the question.**
   For example:

Folder_Name	Table	Columns	Method
Company	employee	salary	median

   Question: What is the median of the employees' salaries?
   Or
   Question: What is the median of salaries for all employees'?

3. **For 20% of the questions, Use synonyms, not the exact columns and tables names.**

   For example:

Folder_Name	Table	Columns	Method
Company	employee	salary	median

   Question: What is the median of the employees' incomes?
   Or
   Question: What is the median of incomes for all workers?

4. **If there is more than one column always write the question columns in the same order as the columns in the table.**
   In the below example, you must write salary in the question before the name.

   For example:

Folder_Name	Table	Columns	Method
Company	employee	Salary * name	scatter

   Question: Show me a scatter plot for the employee's salary and name.
   Or
   Question: Plot a scatter figure for the salary and name of all employees.

5. **Don't write the same patterns for all the questions.**

   For example:
   Don't write
   what is the mean,
   what is the median,
   what is the mode,
   what is the standard deviation,
   what is the normal distribution … etc.

   Be creative like this:
   What is the mean,
   show median of,
   list the most value that appeared,
   give me std of,
   show the normal distribution … etc.

6. **In the Folder names "databases" you can find many <mark>Folders</mark>. Each of them has many <mark>Tables</mark> and each Table has many <mark>columns</mark>.**
You don't need to go to see these tables and columns except if you find a problem in understanding the samples.

For example:

Folder_Name	Table	Columns	Method
baseball_1	team	g * w	weighted_mean

Since you don't know what is g and w columns.

Go to databases → baseball_1 → team
And see what is the meaning of g and w columns

7. **For the following method types, you must define the parameters.**

Method	Parameter1	Parameter2
Weibull distribution	Shape factor	
Chi-square distribution	Degrees of freedom	
t- distribution	Degrees of freedom	
f- distribution	Numerator Degrees of freedom	Denominator Degrees of freedom
Trimmed mean	percent	
percentile	percent	
Confidence interval	percent	

8. **If you cannot write the question for any reason, please write a comment about why you cannot.**

List of following examples

Folder_Name	Table	Columns	Method
activity_1	Student	Age	mean

Example: what is the average age of all students?

Folder_Name	Table	Columns	Method
activity_1	Student	Age	variance

Example: Calculate the var of student ages?

Folder_Name	Table	Columns	Method
activity_1	Student	Grades * Exam_Type	weighted_mean

Example: Show me the weighted mean for the student grades using the exam_type as weights.

Folder_Name	Table	Columns	Method
city_record	temperature	Jan	outlier

Example: Find the outlier temperature degrees in January.

Folder_Name	Table	Columns	Method
city_record	temperature	Jun * Feb	correlation_matrix

Example: using the correlation matrix method show the relations between temperature in June and temperature in February.

Folder_Name	Table	Columns	Method
city_record	temperature	Jun	percentile

Example: What is the 80 percentile of the temperature in the month of June?

Folder_Name	Table	Columns	Method
city_record	temperature	Jun	range

Example: Calculate the difference between the highest and lowest Celsius degrees in June

Folder_Name	Table	Columns	Method
college_2	classroom	capacity	mode

Example: what is the most values appeared in the classroom capacity column.

Folder_Name	Table	Columns	Method
college_2	classroom	capacity	Chi_square distribution.

Example: show the distribution of data for the classroom capacities using chi square distribution with degrees freedom of 2

Folder_Name	Table	Columns	Method
college_2	classroom	capacity	Noraml distribution.

Example: what is the normal distribution for classroom capacities?

Folder_Name	Table	Columns	Method
college_2	student	Gender * Grades	Pie plot

Example: Show me a pie chart for the student's genders using their grades as weight.

Folder_Name	Table	Columns	Method
college_2	student	Gender * Grades	bar plot

Example: Plot a bar graph for the student's genders and grades.

Folder_Name	Table	Columns	Method
college_2	student	Gender * Grades * Age	scatter plot

Example: For the student gender, grade and age show a scatter plot.