

# نمونه سوالات آزمون تجزیه و تحلیل داده های اکتشافی

Exploratory data analysis test

Sample Questions

🕒 سطح آزمون: متوسط

# آزمون تحلیل داده‌های اکتشافی

## Exploratory data analysis

---

آزمون تحلیل داده‌های اکتشافی مهارت شما را در سه حوزه، Feature Engineering، Data Visualization، و Statistical Analysis به چالش می‌کشد.

---

لطفا پیش از شروع آزمون به نکات زیر با دقت، توجه کنید:

- این آزمون از 12 سوال تخصصی تشکیل شده است و نمره منفی ندارد.
- زمان آزمون 20 دقیقه می‌باشد و قابل تمدید نیست.
- این آزمون به زبان انگلیسی طراحی شده است.
- در بعضی از سوالات شاید همه گزینه‌ها صحیح باشند ولی شما باید گزینه‌ای که از بقیه موارد، جامع‌تر و دقیق‌تر می‌باشد را انتخاب کنید.

1- You analyze a dataset where each column signifies the attributes of cars. One column is the car price. Another column is the car color. The car color column has three unique values: "Red," "Black," and "White" (the dataset contains cars of only these three colors).

The screenshot shows the Tableau interface. On the left, a table titled 'Sheet 1' displays data for 'Segment' across 'Category' and 'Sub-Category'. The columns are 'Category', 'Sub-Category', 'Consum.', 'Corpora..', and 'Home Office'. The data is as follows:

Category	Sub-Category	Consum.	Corpora..	Home Office
Furniture	Bookcases	5.91%	4.82%	2.85%
	Chairs	14.88%	14.04%	13.14%
	Furnishings	4.27%	3.54%	3.98%
	Tables	8.60%	10.04%	8.42%
Office Supplies	Appliances	4.55%	5.18%	4.22%
	Art	1.23%	1.22%	1.00%
	Binders	10.17%	7.30%	7.84%
	Envelopes	0.67%	0.84%	0.64%
	Fasteners	0.14%	0.11%	0.13%
	Labels	0.58%	0.58%	0.39%
	Paper	3.13%	3.38%	4.25%
	Storage	8.65%	11.30%	10.14%
	Supplies	2.22%	2.75%	0.35%
Technology	Accessories	7.50%	6.82%	7.47%
	Copiers	6.01%	6.63%	7.65%
	Machines	6.85%	8.54%	11.50%
	Phones	14.63%	12.91%	16.04%

On the right, the 'Table Calculation' dialog box is open, showing '% of Total Sales'. The 'Calculation Type' is 'Percent of Total'. The 'Compute Using' dropdown is set to 'Specific Dimensions'. The 'At the level' dropdown is set to 'Deepest'. The 'Sort order' dropdown is set to 'Specific Dimensions'. The 'Show calculation assistance' checkbox is checked.

How should you reshape the data in the car color column so that you can perform regression using price as a dependent variable?

- Perform a one-hot encoding for all three colors.
- Perform a one-hot encoding for any two colors.
- Replace the color with numeric values (i.e., "Red" as "1", "Black" as "2" and "White" as "3").
- No transformation is required to perform the regression.

2-You are visualizing data for a multi-brand retail store. The store sells three brands of mobile phones and wants to know the percentage contribution of sales for each brand for a given time period.

**Which of the following charts should you use for the visualization?**

- Scatter plot
- Line chart
- Pie chart
- Area chart

3- You have a dataset with 40 independent columns and one dependent column. You decide to drop a few columns to aid stakeholders in determining which attribute they should concentrate on in order to optimize the output.

**Which of the following techniques should you use to decide which columns to drop?**

- A Pearson correlation of all independent columns with the dependent column and drop if the result is less than 0.5.
- A multivariate regression analysis and drop if the p value is greater than 0.05 (Given that  $\alpha = 0.05$ ).
- A Pearson correlation of all independent columns with the dependent column and drop if the result is greater than 0.5.
- A multivariate regression analysis and drop if the p value is less than 0.05 (Given that  $\alpha = 0.05$ ).

**4- True or False:**

In the below table, a chi-square test can be used to determine if there is a statistically significant difference in test results between male and female students.

- ☐ True
- ☐ False