**Attention!** To complete the test tasks, download and open the data array using the link:

[Копия Данные для тестового задания](https://docs.google.com/spreadsheets/d/1TB8gc40MtI4SK0pwo2XOg0o51OHYTSujBjYK7rKlWm4/edit?gid=1687485228#gid=1687485228)

1. In the "Audience Data" tab, information about users who visited our app in November. What is the MAU of the product?

\*MAU (Monthly Active Users) is a metric used to measure user activity for one month. It shows the number of unique users who have interacted with a product, service, or application at least once in the last month.

**7639**168141048216529

Изображение выглядит как текст, Шрифт, снимок экрана

Автоматически созданное описание

2. Using the "Audience Data" tab, calculate what the DAU will be.

\*DAU (Daily Active Users) is a metric that shows the number of unique users who have interacted with a product, application, or service at least once during the day. DAU helps to understand how many users actively use the product every day.

255490560483

Изображение выглядит как текст, снимок экрана, Шрифт, дизайн

Автоматически созданное описание

3. Using the "Audience Data" tab, calculate what the first-day retention rate will be for users who joined the product on November 1st.

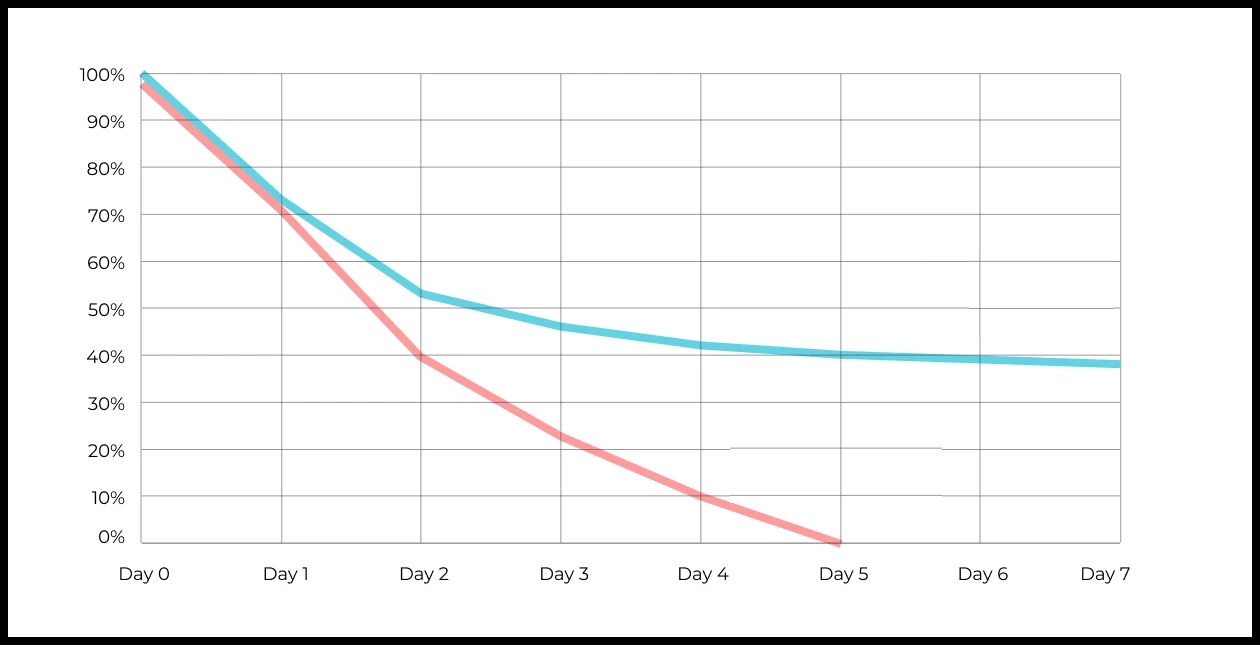
\*Retention is a metric that shows how many users continue to use a product after a certain period following their initial interaction. Retention can be calculated as the percentage of users who returned to the product after a specific time (e.g., after 1 day, 1 week, 1 month) out of the total number of new users.

28,3%**26,6%**38,5%32,7%

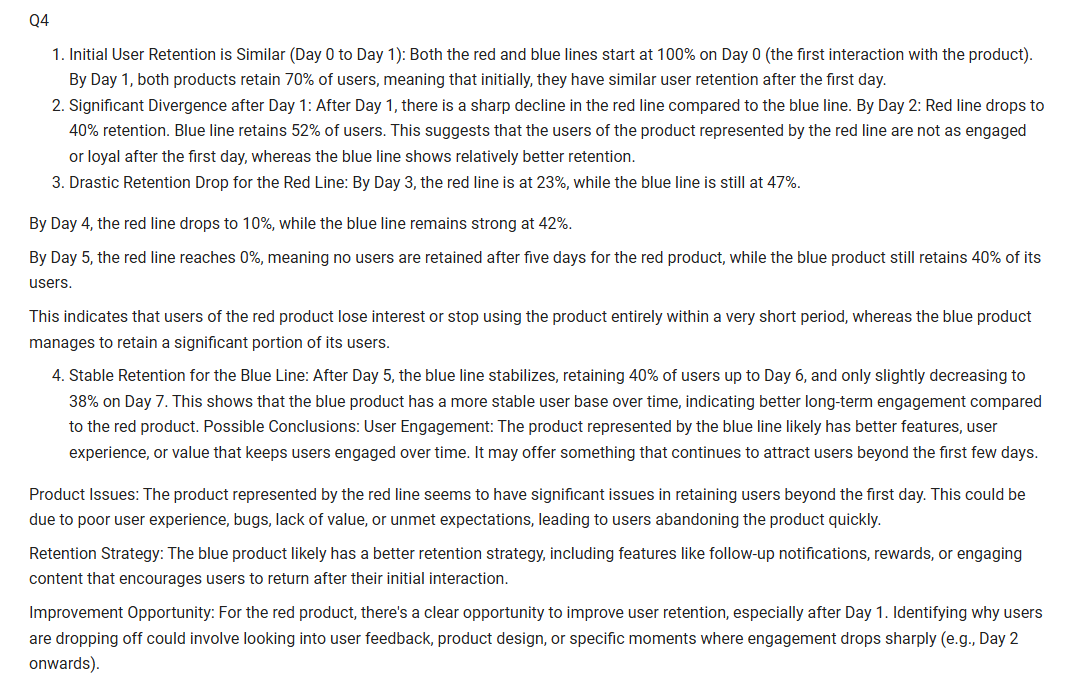
Изображение выглядит как текст, снимок экрана, Шрифт

Автоматически созданное описание

4. On the graph, there are retention curves for two products. What conclusions can be drawn by looking at them?



Your answer:



5. In the "Audience Data" tab, there is information about how many ads each user has viewed (view\_adverts). Do you calculate the user conversion per ad view for November? (in users)

\* User conversion is a metric that shows what percentage of users completed a target action relative to the total number of users. In the context of websites, this can be an action such as viewing an ad or clicking on an advertising banner.

41,8%54,7%**46,3%**39%

Изображение выглядит как текст, снимок экрана, Шрифт, дизайн

Автоматически созданное описание

6. Using the information from the Audience Data tab, calculate the average number of ads viewed per user in November

4,96,25,3**2,9**

Изображение выглядит как текст, снимок экрана, Шрифт, дизайн

Автоматически созданное описание

7. We conducted a survey among 2,000 users. Of them, 500 are "critics," 1,200 are "promoters," and 300 are "neutrals." Calculate the NPS.

\*NPS (Net Promoter Score) — is a metric that measures user loyalty towards a company or product and categorizes them into three groups: Promoters, Passives, and Detractors. NPS is calculated as (% of promoters - % of detractors).

30%43%40%**35%**  
Изображение выглядит как текст, снимок экрана, Шрифт

Автоматически созданное описание

8. In the "AB Test Data" tab, there are results of three unrelated AB tests for ARPU (total revenue/total number of users).

1. **Look at the test results and interpret them.**
2. **Write down the p-values you obtained.**
3. **Prepare conclusions and recommendations.**

**Data Columns:**

* **experiment\_num**: Experiment number
* **experiment\_group**: Group the user was assigned to
* **user\_id**: User ID
* **revenue**: Revenue generated by the user through purchasing a paid promotion service.

Your answer:

Изображение выглядит как текст, снимок экрана, Шрифт

Автоматически созданное описаниеИзображение выглядит как текст, снимок экрана, Шрифт, число

Автоматически созданное описаниеИзображение выглядит как текст, снимок экрана, Шрифт, документ

Автоматически созданное описание

9. Calculate the average income per user based on the dataset with the listers

121.2**156.4**70.930.7the average is not applicable here  
Изображение выглядит как текст, снимок экрана, Шрифт, число

Автоматически созданное описание

10. Based on the dataset with the listers, calculate the median age of the user

27,42**28**27,9327the median is not applicable here

Изображение выглядит как текст, снимок экрана, Шрифт, число

Автоматически созданное описание

11. Which chart is best suited to display the spread of prices for goods in different stores?

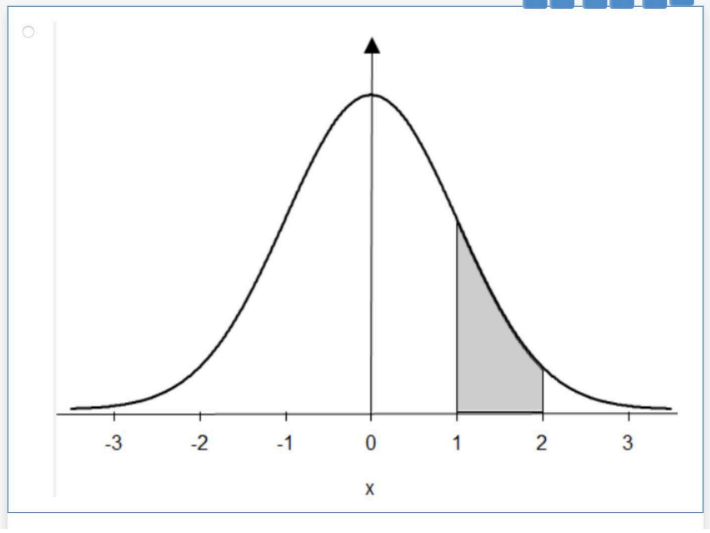
\*There may be several possible answers.

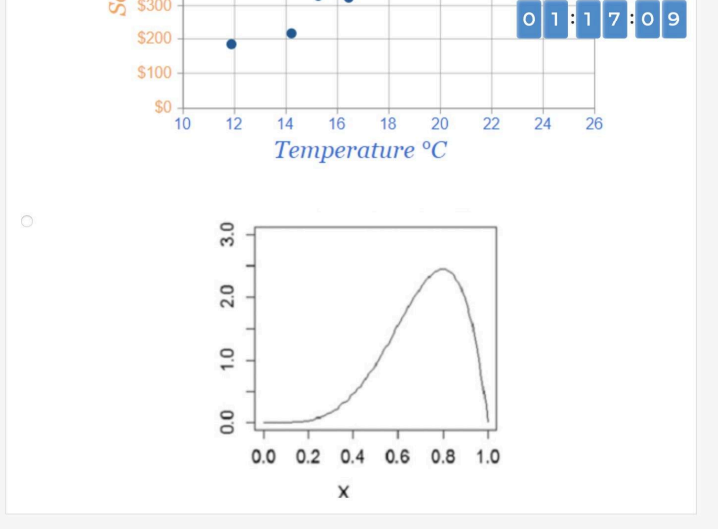
Line chartPie chart**Box with whiskers (box plot)**

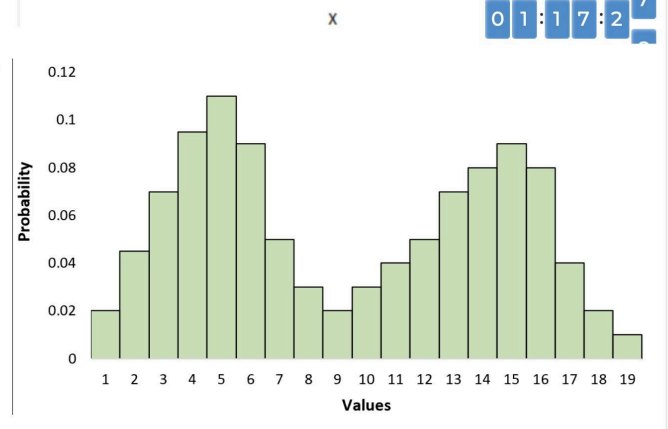
The histogram  
Изображение выглядит как текст, снимок экрана, Шрифт, алгебра

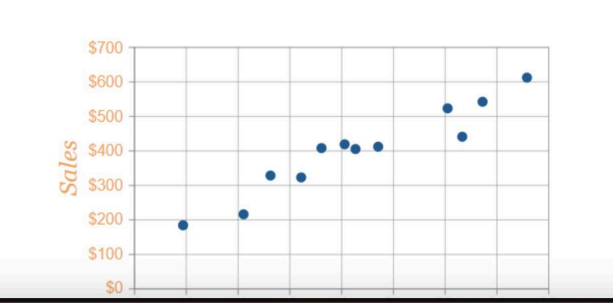
Автоматически созданное описание

12.What is the bimodal distribution graph?

№1

№2

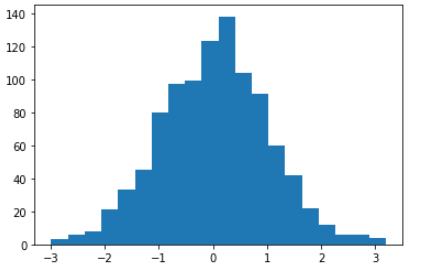
**№3**

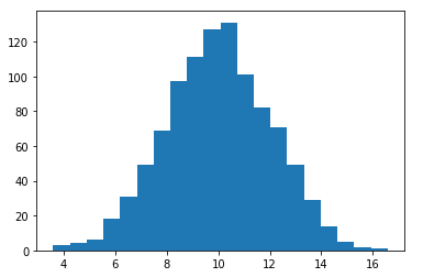
№4

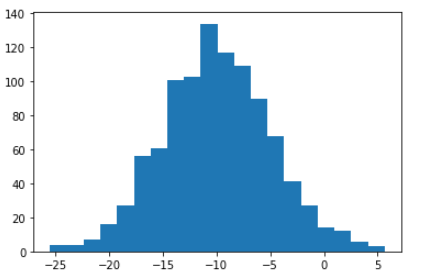
The **bimodal distribution** graph is the one that shows two distinct peaks or modes. In the images you provided, the graph labeled **№3** (the green bar chart) is the bimodal distribution. This histogram displays two separate clusters of data with peaks around different values, which is characteristic of a bimodal distribution.

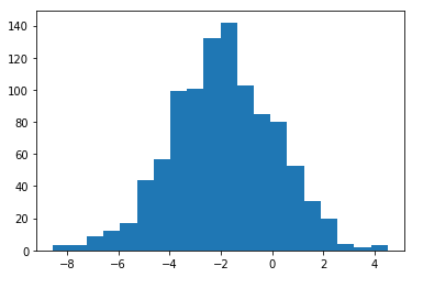
**Answer: №3**

13. Which random variable has the largest data variance according to the following distribution density graph?

№1

№2

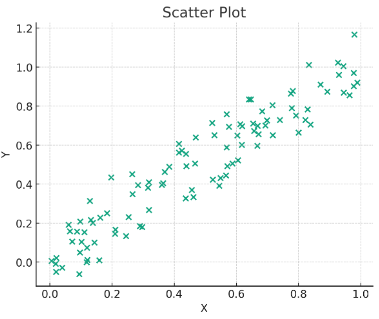
**№3**

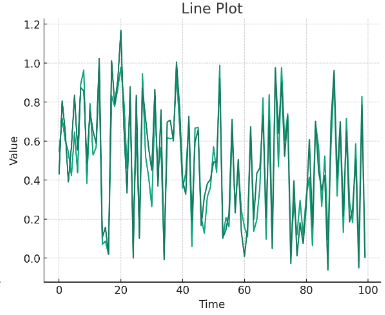
№4

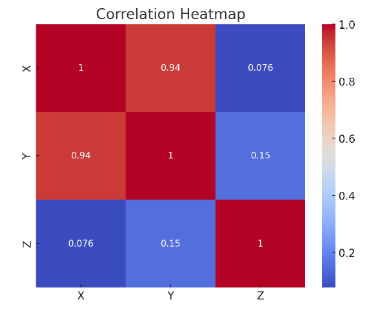
**№1, №2, №3, and №4** have similar shapes, but the graph labeled **№3** has the widest spread, extending from approximately -30 to 5. This indicates a larger range of values and, consequently, a higher variance.

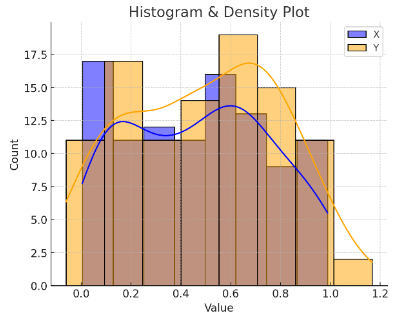
**Answer**: The random variable in **№3** has the largest data variance.

14. On which graph can the correlation be calculated?  
\**There may be several possible answers.*









**Scatter Plot**: The scatter plot (first graph) shows the relationship between two variables (X and Y). Correlation can be calculated here by assessing how closely the points follow a linear pattern.

**Correlation Heatmap**: The correlation heatmap (second graph) explicitly displays the correlation coefficients between variables X, Y, and Z. This visualization directly presents the correlation values.

**Answer**: **Scatter Plot** and **Correlation Heatmap**.

15. What does it mean if, when testing hypotheses, we got p-value = 0.05?

This means that there is no statistically significant difference between the groups

**There is a 5% chance of accidentally getting this or an even more extreme result if the null hypothesis is correct**

This means that the results of the experiment are 95% accurate

This suggests that the alternative hypothesis is correct with a probability of 95%

16. Which method is most suitable for testing the hypothesis of equality of the average of two samples from a normal distribution?

![Изображение выглядит как снимок экрана, пиксель, дизайн

Автоматически созданное описание](data:image/png;base64,iVBORw0KGgoAAAANSUhEUgAAADcAAAAwCAMAAACCABLMAAAAAXNSR0IArs4c6QAAAARnQU1BAACxjwv8YQUAAAAYUExURSYmJqCgoGlpaf///z4+PuPj45WVlQAAAIGrJ1gAAAAIdFJOU/////////8A3oO9WQAAAAlwSFlzAAAh1QAAIdUBBJy0nQAAAHZJREFUSEvt0rEOgDAIRVGw1v//ZKG8amKsIkNduCNwhialLJscn2HiC0bDxJOeL9YnGXNmimUWm+dibhUF06u1FGyHtZfhvqfuDYbddlWaSOxHpbtvupv8X6L/WmKu0mGcKu6IOmjGreLOJMLEF4yGSZb9GdEOHgwF94hF65sAAAAASUVORK5CYII=)**t-test**Chi-square testAnalysis of Variance (ANOVA)Pearson Correlation

17. How to interpret quartiles in the distribution of user income?

They show the maximum and minimum income

**Divide the data into four equal parts**

Indicate the most common income

Graph of the density distribution of matter in the universe

18. The following results were obtained. Colleagues ask you to confirm them and make a final conclusion on the experiment.

* Option A (control group) — 100,047,501 visitors, 1003 payments.
* Option B (test group) — 100,001,055 visitors, 1099 payments.

What recommendations would you make based on this data?

**Your answer:**

1. **Conversion Rates**:
   * **Control Group (Option A)**: 1003 payments out of 100,047,501 visitors (conversion rate ~0.001002%).
   * **Test Group (Option B)**: 1099 payments out of 100,001,055 visitors (conversion rate ~0.001099%).

The test group has a slightly higher conversion rate compared to the control group.

1. **Recommendation**:
   * **If the increase is statistically significant**: This suggests the changes in the test group had a positive effect, and we should consider rolling out these changes to all users.
   * **If not statistically significant**: The difference might be due to random variation, and further testing could be needed to confirm any real impact.

In summary, if statistical testing confirms the difference, proceed with the changes; otherwise, gather more data.