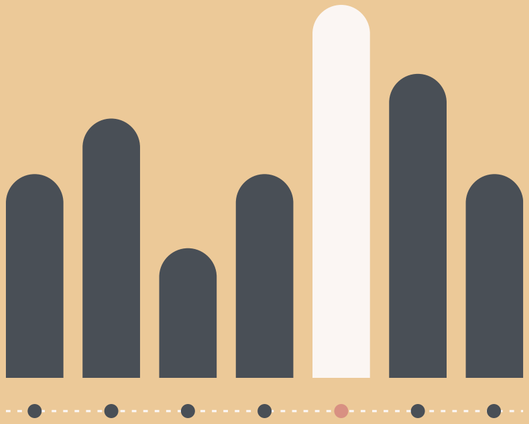
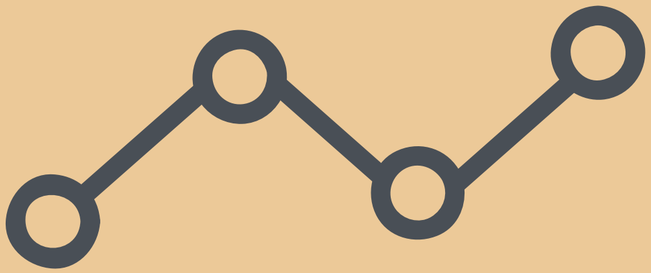
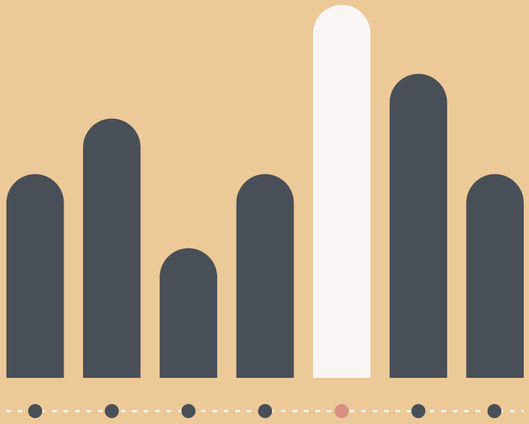
NAME THE GRAPH

Statistics – Data Displays Quiz







TEAM

Rawnok Jahan Riddi
>0242220005101830

Md Raduan Ahamed
>0242220005101839

Md Sajib Hossen
>0242220005101826

Md Hafijur Rahman
>0242220005101620

What is Hypothesis testing?

It involves evaluating whether a particular assumption (or hypothesis) about a population parameter is consistent with the observed data.

What is Hypothesis testing?

Hypothesis testing is a statistical method used to make decisions or inferences about a population based on sample data.

Hypotheseis Test

Step-1: Formulate hypotheses

Two-sample t-test:

- Research Question: Is there a significant difference in the mean age between individuals with lung cancer and those without?
- Null Hypothesis (Ho): There is no significant difference in the mean age between the two groups.
- Alternative Hypothesis (H1): There is a significant difference in the mean age between the two groups.

Hypotheseis Test

Step-1: Formulate hypotheses

Chi-square test:

- **Research Question**: Is there an association between smoking status and lung cancer?
- **Null Hypothesis (Ho)**: Smoking status is not associated with lung cancer.
- **Alternative Hypothesis (H1)**: Smoking status is associated with lung cancer.

Hypotheseis Test

Step 2: Select and Apply Relevant Hypothesis Tests

- Two-sample t-test for comparing means of scores between the two groups.
- Chi-square test for testing the association between smoking status and lung cancer.

Two-sample T-Test

```
import pandas as pd
from scipy.stats import ttest_ind, chi2_contingency

df = pd.read_csv('/mnt/data/dataset-hafiz.csv')

age_lung_cancer_yes = df[df['LUNG_CANCER'] ==
'YES']['AGE']
age_lung_cancer_no = df[df['LUNG_CANCER'] ==
'NO']['AGE']

t_stat, p_value = ttest_ind(age_lung_cancer_yes,
age_lung_cancer_no, equal_var=False)

print(f"T-statistic: {t_stat:.4f}, p-value: {p_value:.4f}")
```

Output

```
T-statistic: -1.9573,
p-value: 0.0504
Fail to reject the
null hypothesis
```

Two-sample T-Test

```
alpha = 0.05  # Significance level
if p_value <= alpha:
    print("Reject the null hypothesis.")
else:
    print("Fail to reject the null
hypothesis.")
```

Output

```
T-statistic: -1.9573,
p-value: 0.0504
Fail to reject the
null hypothesis.
```

THANK YOU