1. You work at an e-commerce company that sells three goods: widgets, doodads, and fizzbangs. The head of advertising asks you which they should feature in their new advertising campaign. You have data on individual visitors' sessions ([activity on a website](https://en.wikipedia.org/wiki/Session_%28web_analytics%29), [pageviews](https://en.wikipedia.org/wiki/Page_view), and purchases), as well as whether or not those users [converted](https://en.wikipedia.org/wiki/Conversion_marketing) from an advertisement for that session. You also have the cost and price information for the goods.

|  |  |  |  |
| --- | --- | --- | --- |
|  | Widgets | Doodads | Fizzbangs |
| Converted | Cost: $10  # of visitor sessions: 200  # of Pageviews: 150 | Cost: $25  # of visitor sessions: 300  # of Pageviews: 250 | Cost: $65  # of visitor sessions: 350  # of Pageviews: 275 |
| Didn’t convert | Cost: $45  # of visitor sessions: 125  # of Pageviews: 110 | Cost: $35  # of visitor sessions: 75  # of Pageviews: 100 | Cost: $20  # of visitor sessions: 135  # of Pageviews: 115 |

**Testable**: Put the information into a matrix that shows rows for converted/non-converted for each type of good.

**Translate to question**: Which good has a higher conversion rate based on the number of visitors per sessions.

1. You work at a web design company that offers to build websites for clients. Signups have slowed, and you are tasked with finding out why. The [onboarding funnel](https://en.wikipedia.org/wiki/Funnel_analysis) has three steps: email and password signup, plan choice, and payment. On a user level you have information on what steps they have completed as well as timestamps for all of those events for the past 3 years. You also have information on [marketing spend](https://en.wikipedia.org/wiki/Marketing_spending) on a weekly level.

|  |  |  |  |
| --- | --- | --- | --- |
|  | Email | Signup | Plan choice |
| Completed | Cost: $10  # of emails: 200  # of Pageviews: 150 | Cost: $25  # signups: 300  # of Pageviews: 250 | Cost: $65  # of plan choice: 350  # of Pageviews: 275 |
| Not completed | Cost: $45  # of emails: 125  # of Pageviews: 110 | Cost: $35  #of emails: 200  # of Pageviews: 100 | Cost: $20  # of plan choice: 135  # of Pageviews: 115 |

**Testable**: Put the information into a matrix that shows rows for completed/not-completed for onboard funnel process.

**Translate to question**: How many users completed the entire onboard funnel process? How many users only completed the sign-up portion?

1. You work at a hotel website and currently the website ranks search results by price. For simplicity's sake, let's say it's a website for one city with 100 hotels. You are tasked with proposing a better ranking system. You have session information, price information for the hotels, and whether each hotel is currently available.

|  |  |  |  |
| --- | --- | --- | --- |
|  | Best | Average | Bad |
| Hotel available | Cost: $10  # of ratings: 200  # of sessions: 150 | Cost: $25  # of ratings: 300  # of sessions: 250 | Cost: $65  # of ratings: 350  # of sessions: 275 |
| Hotel not available | Cost: $45  # of ratings: 125  # of sessions: 110 | Cost: $35  # of ratings: 200  # of sessions: 100 | Cost: $20  # of ratings: 135  # of sessions: 115 |

**Testable**: Put the information into a matrix that shows rows for hotel available /non-available hotel for each type of rating.

**Translate to question**: How many hotels have an average rating with a price >100? How many hotels are not available with a best rating? How many hotels are rated bad with a high price?

1. You work at a social network, and the management is worried about [churn](https://en.wikipedia.org/wiki/Churn_rate) (users stopping using the product). You are tasked with finding out if their churn is atypical. You have three years of data for users with an entry for every time they've logged in, including the timestamp and length of session.

|  |  |  |
| --- | --- | --- |
|  | Using product | Stop using product |
| Typical | Cost: $10  # of loggin times: 200  # length of sessions: 150 | Cost: $25  # of loggin times: 300  # length of sessions: 250 |
| Not typical | Cost: $45  # of loggin times: 125  # length of sessions: 110 | Cost: $35  # of loggin times: 200  # length of sessions: 100 |

**Testable**: Put the information into a matrix that shows rows for typical and not typical for each type of product.

**Translate to question**: For a typical user what is the maxium length of session time they are using the product?