# 2024/05/22

## Ask Questions to Make Data-Driven Decisions

### Module 2: Data trials and triumphs

#### Data-driven decisions

data-driven decision-making means using facts to guide business strategy.

Data is used to arrive at a decision. This approach is limited by the quantity and quality of readily-available data.

**Potential dangers** of relying entirely on data-driven decision-making can include **overreliance on historical data**, **a tendency to ignore qualitative insights**, and **potential biases in data collection and analysis**

**Example: A/B testing, compared the influence of A and B**

#### Data-inspired decisions

They create space for people using data to consider a broader range of ideas: drawing on comparisons to related concepts, giving weight to feelings and experiences, and considering other qualities that may be more difficult to measure.

Example: CSAT score

#### Qualitative and quantitative data in business

CASE STUDY

Now, take a closer look at the data types and data collection tools. In this scenario, you are a data analyst for a chain of movie theaters. Your manager wants you to track trends in:

* **Movie attendance over time**

You look at attendance over the last 3 months. But, because the last 3 months didn’t include a major holiday, you decide it is better to look at a full year’s worth of data. (results: the move attendance in the month with holidays higher )

you also decide that you will resume the analysis again in a few months after the theater increases ticket prices for evening showtimes.

* **Profitability of the concession stand**

The historical data shows that while the concession stand was profitable, profit margins were razor thin at less than 5%.

You suggest an online survey of customers so they can comment on the food at the concession stand. This will enable you to gather even more quantitative data to revamp the menu and potentially increase profits.

* **Evening audience preferences**

Your analysis of the historical data shows that the 7:30 PM showtime was the most popular and had the greatest attendance, followed by the 7:15 PM and 9:00 PM showtimes.

Evening movie-goers are the largest source of revenue for the theater. Therefore, you also decide to include a question in your online survey to gain more insight.

#### Qualitative data for all three trends plus ticket pricing

Since you know that the theater is planning to raise ticket prices for evening showtimes in a few months, you will also include a question in the survey to get an idea of customers’ price sensitivity.

Pivot Table: 数据透视图

Revenue: number of sales \* sales price

ROI: return on investment 投资回报率

#### Tools for visualizing data

Spreadsheets

creating static charts and graphs

clean, sort, and filter data.

Simple visualization

Tableau

Advanced data visualization and analytics capabilities

easy to create charts, graphs, and dashboards in a drag-and-drop interface.

Tableau can handle more data and larger datasets than many other tools and offers real-time data availability.

Drag-and-drop Interface. 拖放界面！

Versatility 多功能性的

找工作当然烦，但是通过自己的能力找到一份自己想干的事情，对于自己来说不是一件很有成就感的事情吗！这件事情就算有别人的意见，但是主体是你自己，你究竟想要怎么样的人生和工作！

#### Create a dashboard

**1. Identify the stakeholders who need to see the data and how they will use it**

**2. Design the dashboard (what should be displayed)**

Labels, short description, show the most important information at the top

**3. Create mockups if desired**

**4. Select the visualizations**

show a change in values over time, line charts or bar graphs might be the best choice

show how each part contributes to the whole amount being reported, a pie or donut chart is probably a better choice.

**5.** **Create filters as needed**

#### Big and small data

| **Small data** | **Big data** |
| --- | --- |
| Describes a dataset made up of specific metrics over a short, well-defined time period | Describes large, less-specific datasets that cover a long time period |
| Usually organized and analyzed in spreadsheets | Usually kept in a database and queried |
| Likely to be used by small and midsize businesses | Likely to be used by large organizations |
| Simple to collect, store, manage, sort, and visually represent | Takes a lot of effort to collect, store, manage, sort, and visually represent |
| Usually already a manageable size for analysis | Usually needs to be broken into smaller pieces in order to be organized and analyzed effectively for decision-making |

#### The three (or four) V words for big data

| **Volume** | **Variety** | **Velocity** | **Veracity** |
| --- | --- | --- | --- |
| The amount of data | The different kinds of data | How fast the data can be processed | The quality and reliability of the data |

# 2024/05/23

# Spreadsheets and the data life cycle

the data life cycle: **plan, capture, manage, analyze, archive,** and **destroy**.

### Plan:

developing organizational standard, including formatting your cells, the headings you choose to highlight, the color scheme, and the way you order your data points.

Results: you will improve communication, ensure consistency, and help people be more efficient with their time.

图示

描述已自动生成

### Capture:

connecting spreadsheets to other data sources

automatically be updated in the spreadsheet.

### Manage:

storing, organizing, filtering, and updating information

Spreadsheets also let you decide who can access the data, how the information is shared, and how to keep your data safe and secure.

### Analyze:

**formulas** to aggregate data or create reports, and **pivot table**s for clear, easy-to-understand visuals.

### **Archive:**

store historical data before it gets updated.

### Destroy: