

Data Analytics Project

6. The Data Selection Guide for Analysis



Aagam Deolasi
@Aagam Deolasi



The Data Selection Journey

If you remember, **Asking Questions** is the first and foremost phase of *Data Analytics Life Cycle (DALC)**.

Selecting relevant data for your analysis includes *determining the type(s) of data* needed and *finding a suitable source*. It's crucial to focus on data directly linked to your business question *to ensure validity and reliability*.

Now, let's explore some of the **key questions** in the **data selection process**.

* I have covered DALC in Part 2 of the series.



2

Data Source Selection

A question (we can ask):

What ***data points*** are ***necessary to inform*** your analysis?

Example:

If you're ***analyzing customer satisfaction***, necessary data points might include customer feedback scores, purchase history, and demographics.



Aagam Deolasi
@Aagam Deolasi

3

Data Accessibility

A question (we can ask):

Do I ***already have access to this data***, or must I ***find a dataset from another source***?

Example:

I have access to internal sales data, but for ***market trends***, I may need to explore external databases or collaborate with industry partners.



Aagam Deolasi
@Aagam Deolasi

4

Source Reliability

A question (we can ask):

Where are ***reliable and verifiable sources*** of this data?

Example:

For ***economic indicators***, reliable sources could include government reports or reputable financial institutions known for accurate and transparent data reporting.



5

Data Collection Frequency

A question (we can ask):

How often is the relevant ***data collected and updated?***

Example:

Daily sales data might be crucial for retail analysis, while ***census data*** might be updated less frequently but still provide valuable demographic insights.



Aagam Deolasi
@Aagam Deolasi

Data Licensing and Cost

A question (we can ask):

How *is the data licensed* for use, and *is there a cost*?

Example:

Industry reports may require a subscription or a one-time purchase, while some *open datasets* may have Creative Commons licenses allowing free use with attribution.



Data Format Compatibility

A question (we can ask):

Is the data in a format that I can use, ***or convert to use***, with my tools?

Example:

If data is in a ***CSV format***, it's easily importable into most analysis tools; however, if it's in a ***proprietary format***, you may need to assess conversion possibilities.



THANK YOU!!! FOR YOUR SUPPORT! For Now...

Keep Learning, Keep Sharing & Keep Following
Aagam Deolasi.



Aagam Deolasi
@Aagam Deolasi

