Purpose of EDA

the questions to be answered serves as a guide in EDA

Polish the Questions
Check if the questions to be answered are valid or well stated; If not, modify them or come up with new ones

Validate Data I/O Methods
Check and validate the methods to load and save the datasets

Is the Dataset Good Enough for the Problem?

Are the features/variables required for the project included?
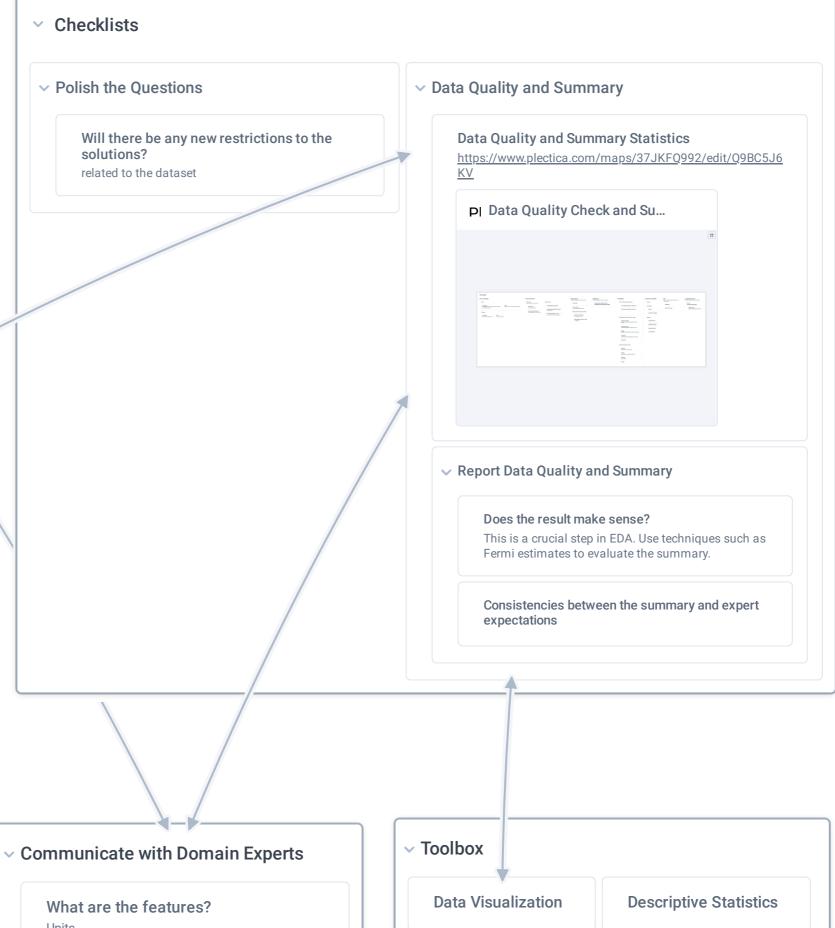If not, what other data should be included.

What is the General Quality of the Dataset

Can one answer the questions semi-quantitatively using the data?

Retrieve Domain Knowledge and Anomalies

Propose the Next Steps
Feasibility of the project

Checklists

Polish the Questions

Will there be any new restrictions to the solutions?
related to the dataset

Data Quality and Summary

Data Quality and Summary Statistics
https://www.plectica.com/maps/37JKFQ992/edit/Q9BC5J6KV

P| Data Quality Check and Su...

Report Data Quality and Summary

Does the result make sense?
This is a crucial step in EDA. Use techniques such as Fermi estimates to evaluate the summary.

Consistencies between the summary and expert expectations

Communicate with Domain Experts

What are the features?
Units

Do the results from EDA make sense to the experts?

What do the experts want to know from the data?

Toolbox

Data Visualization

Descriptive Statistics

EDA