Columns and DataType Not Explicitly Set

Context

In Pandas, all columns are selected by default when a `DataFrame` is imported from a file or other sources. The data type for each column is defined based on the default `dtype` conversion.

Problem

- If columns are not explicitly selected, it becomes unclear what to expect in the data schema for downstream processes.
- If data types are not set explicitly, the process may silently move to the next step with unexpected inputs, potentially leading to errors later.

Therefore, every data loading operation **should contain the name of the column AND the datatype of each column.**

Existing Stage	Effect
Data importing and **Cleaning**.	Readability

Solution

It is recommended to specify **columns name AND datatypes when loading data.**

Example

Python

```
### Pandas Column Selection
import pandas as pd
df = pd.read_csv('data.csv')
+ df = df[['col1', 'col2', 'col3']]
### Pandas Set DataType
import pandas as pd
- df = pd.read_csv('data.csv')

+ df = pd.read_csv('data.csv', dtype={'col1': 'str', 'col2':

'int', 'col3': 'float'})
```