



Analytics Canvas Tutorial:

Data Preparation Considerations



Data Preparation Considerations

A company's information is one of their most important assets [1]. Data sources can be found throughout an organization (internal) and beyond its borders (external), in both structured and unstructured forms. Internal data sources in the form of spreadsheet files, databases, data warehouses and data marts, store a wealth of useful data. External databases and online sources are also valuable resources that can be combined with internal sources to produce business intelligence.

Data quality is a must for effective data analysis. Without ensuring that data is fit for the purpose, it would be impossible to trust the insights derived from it. The most important data quality requirements are:

- Data completeness (no noisy or missing data);
- Data consistency (no inconsistent variables and variable names);
- Data reliability (no sampled data).

Conclusion

Thank you for reading this tutorial. We invite you to continue with the tutorial training to learn more about using Analytics Canvas.

References

[1] Gartner, <http://www.gartner.com/smarterwithgartner/why-and-how-to-value-your-information-as-an-asset/>