



## WHY DATA PREPARATION IS AN IMPORTANT TASK

By Olga Papadimitriou – Sales Administrator

Data cleansing, transformation, and restructuring are all steps in the process of data preparation for analysis, business intelligence, and visualization. Putting data in context in the age of big data is frequently a time-consuming effort for data engineers or users, but it is crucial. Through this process, mistakes and bias brought on by bad data quality are removed and data are transformed into insights. A crucial phase in the data management process is data preparation, which can assist guarantee that data is correct, consistent, and prepared for modelling.

Data preparation tasks such as merging multiple data sources, splitting datasets, and transforming data can be complicated as it can involve a variety of tasks, like data cleaning (removing inaccurate or missing data), data transformation (changing data from one format to another), and data restructuring (aggregating data or creating new features). Although data preparation might be time-consuming, it is necessary for the development of precise predictive models.

### The importance of data preparation

Preparing data takes up much of the time that data scientists have. Many data scientists feel that data preparation is the least enjoyable aspect of their work because of the amount of time they must spend on tedious activities, but reliable insights can only come from well-prepared data. Here are some of the main justifications for the significance of data preparation:

- Delivers accurate analytics application findings
- Encourages wiser decision-making
- Decreases the cost of data management and analytics
- Prevents redundant work
- Increases the ROI of BI and analytics projects

Each business and engineer may have a different method for preparing data. However, the data preparation procedure consists of six basic steps:

- 1) Data gathering
- 2) Data profiling and discovery
- 3) Data cleansing
- 4) Data structuring



- 5) Data enrichment and transformation
- 6) Verifying and publishing data

### Data preparation tips for businesses

When preparing data for diverse business use cases, keep in mind these six suggestions:

1. **Get ready for preparation.** Make a schedule for who will finish which preparation tasks, on what schedule, and for what business objectives.
2. **Don't pretend that the data is flawless.** Make sure you let stakeholders know if there are any data restrictions so you can adjust expectations as soon as feasible.
3. **While tools can be helpful, humans are still crucial.** The inclusion of a human (with the necessary abilities) guarantees data quality, more accurate results, and gives the data context.
4. **To comprehend the distribution of your data, use hypothesis testing.** You can perform several tests to gain a deeper understanding of the data and better prepare it for use.
5. **Set data priorities based on your use case.** To prioritize data sources for a particular model, human judgment is necessary. The data preparation procedure can be streamlined by selecting which data sources will be used first.
6. **Consider data storage carefully.** The pain of data preparation can be greatly reduced by standardizing data formats when data is being processed.

### Data preparation tools

Data preparation is a time-consuming activity that many individuals would choose to completely avoid. However, there are various tools for data preparation that can make the procedure easier, automated, and faster. One of the leaders in this market is Microsoft which is named a Leader in the March 2022 Gartner® Magic Quadrant™ for Analytics and Business Intelligence Platforms. Microsoft is again positioned furthest in completeness of vision and highest in ability to execute for the fourth year running.

### What is Microsoft Power BI?

You can connect to and display any data with Power BI, which offers a single, scalable platform for self-service and corporate Business Intelligence (BI). This platform is simple to use and aids in providing you with deeper data insight. Data analysis and decision-making are combined with Power BI. It enables you to connect to, model, and visualize your data, create reports that are tailored to your KPIs and brand, and get quick, AI-powered responses to your business queries. Additionally, by connecting to all your data sources on a large scale and preserving data accuracy, consistency, and



security, you can maximize the return on your big data investments while analysing, sharing, and promoting insights within your business. Finally, you can easily collaborate on reports, work on the same data, and share insights across well-known Microsoft Office programs like Microsoft Teams and Excel – enabling everyone in your business to make data-driven choices that lead to strategic actions right away.

Utilizing dependable Microsoft technology, Power BI allows you to maximize your business intelligence:

- End-to-end data protection.
- Utilize Power BI with Azure and Office to get the most out of your data and technologies.
- Have a large collection of data connections (with a library of 500+ free connectors still expanding), so that everyone may make data-driven decisions. Direct access to hundreds of local and remote data sources, including Dynamics 365, Azure SQL Database, Salesforce, Excel, and SharePoint.

With Power BI Desktop, you can generate and store an endless number of interactive reports, share them with others, and collaborate on them. This makes Microsoft Power BI ideal for self-service and enterprise business intelligence. Additionally, you can view and interact on reports and visualizations anywhere with the Power BI mobile app for Android, iOS, and Windows Mobile.

As a Microsoft Gold Partner in Cyprus, IBSCY Ltd provides businesses with a variety of IT services and solutions that help them evolve efficiently and without difficulties at a time when digital business upgrades require significant adjustments to their conventional mode of operation.