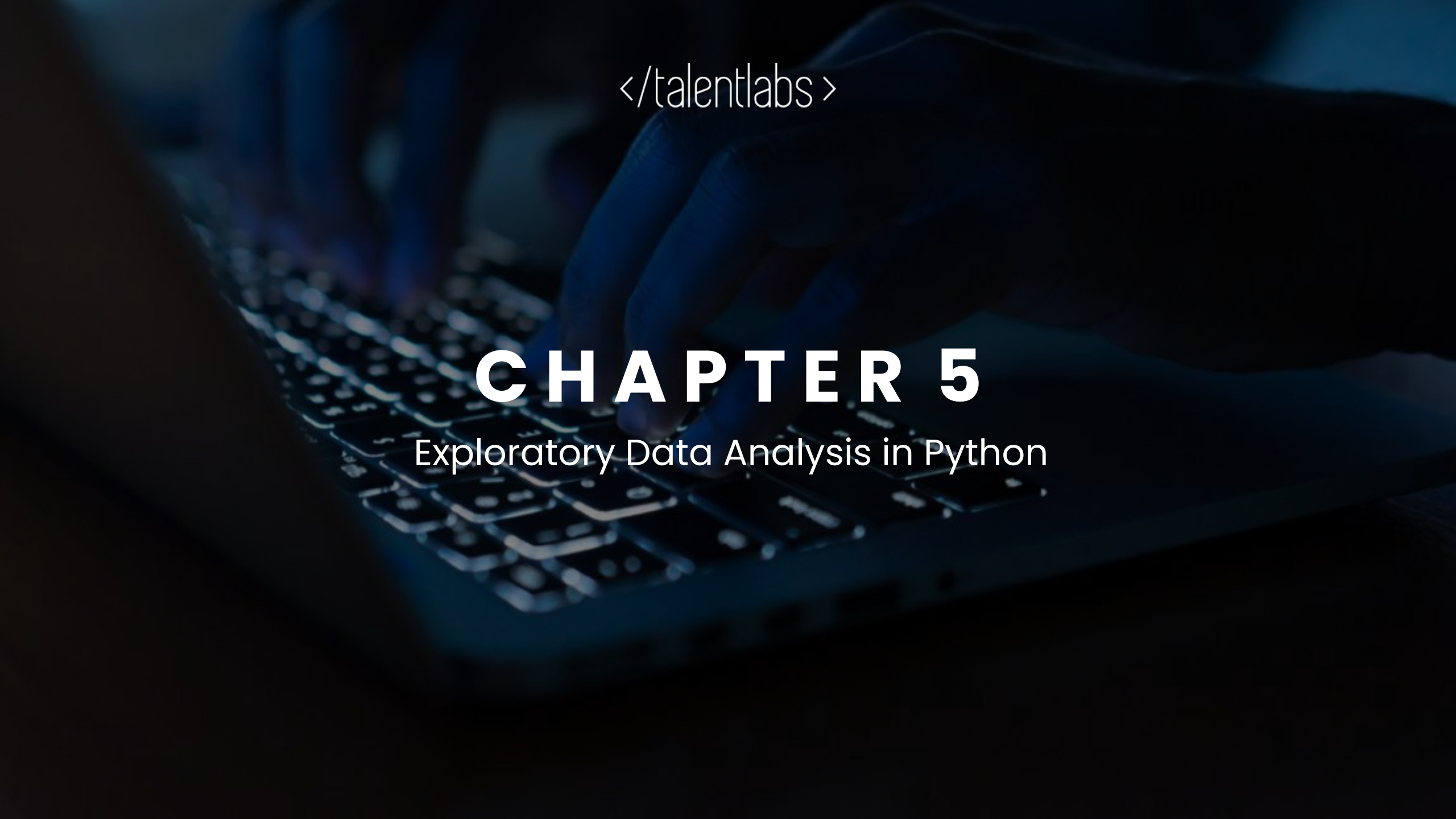

</talentlabs>
CHAPTER 5
Exploratory Data Analysis in Python

Objectives

- Understanding Exploratory Data Analysis, the Data Analytics pipeline and how to generate Business Value.
- Hands on experience of Graphic systems in Python - Matplotlib, Seaborn and Pandas plotting to enable an analyst to visualize data with the right toolset.
- Diving deep into analyzing structure of the data and extracting insights.
- Interpreting different kinds of charts and plots.
- Carrying out basic statistical analysis and understanding Distribution of data.

Chapter Outline

- Introduction to Analytics and Exploratory Data Analysis
- Plotting Systems in Python
- Univariate Analysis
- Statistical Analysis of Distributions
- Bivariate Analysis
- Multivariate Analysis
- End to End Exploratory Data Analysis in Python
- Conclusion

</talentlabs>

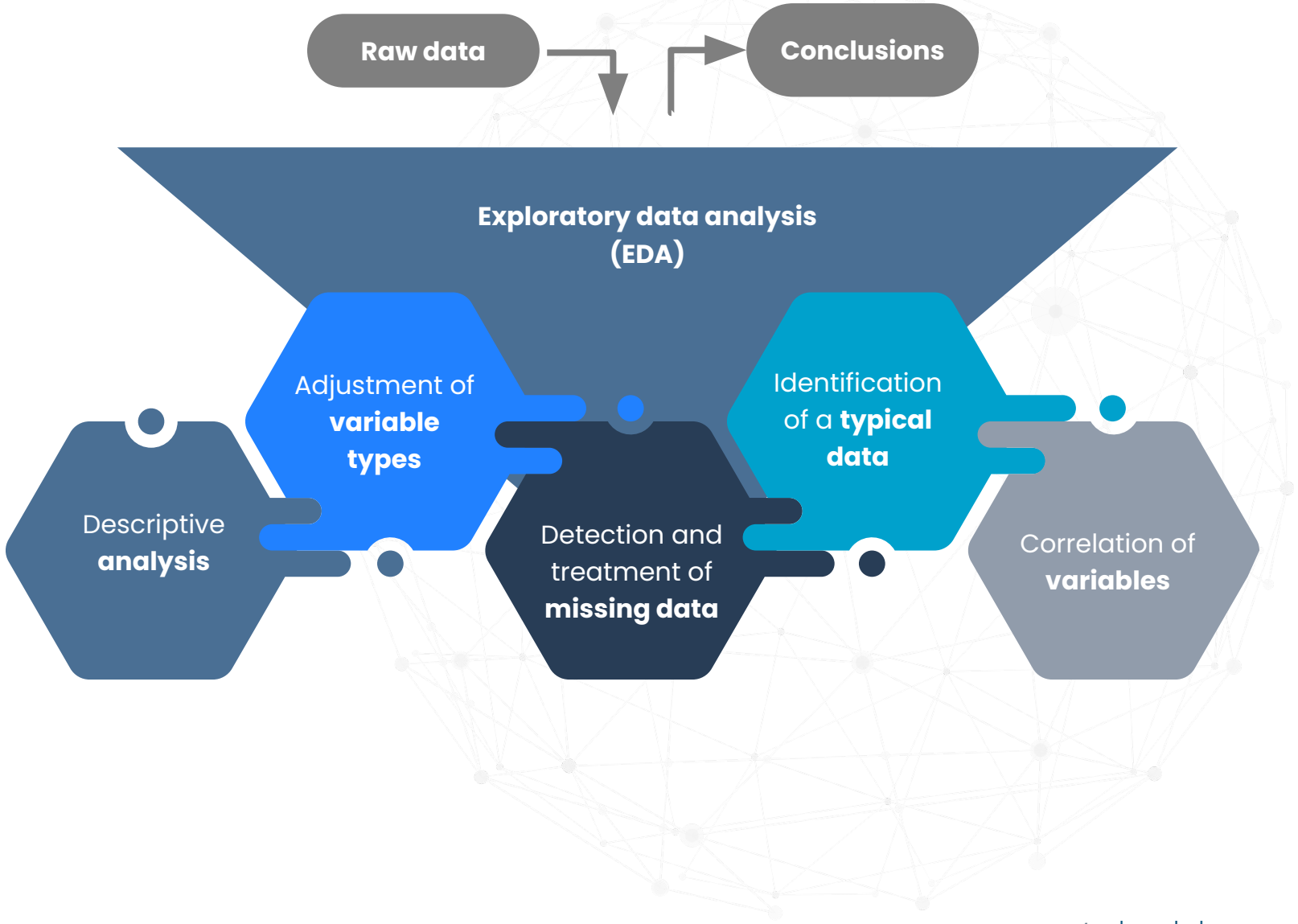
Agenda

- What is Exploratory Data Analysis?
- What is the Exploratory Data Analysis Pipeline?
- How is EDA helpful in taking this data and generating business value?

</talentlabs>

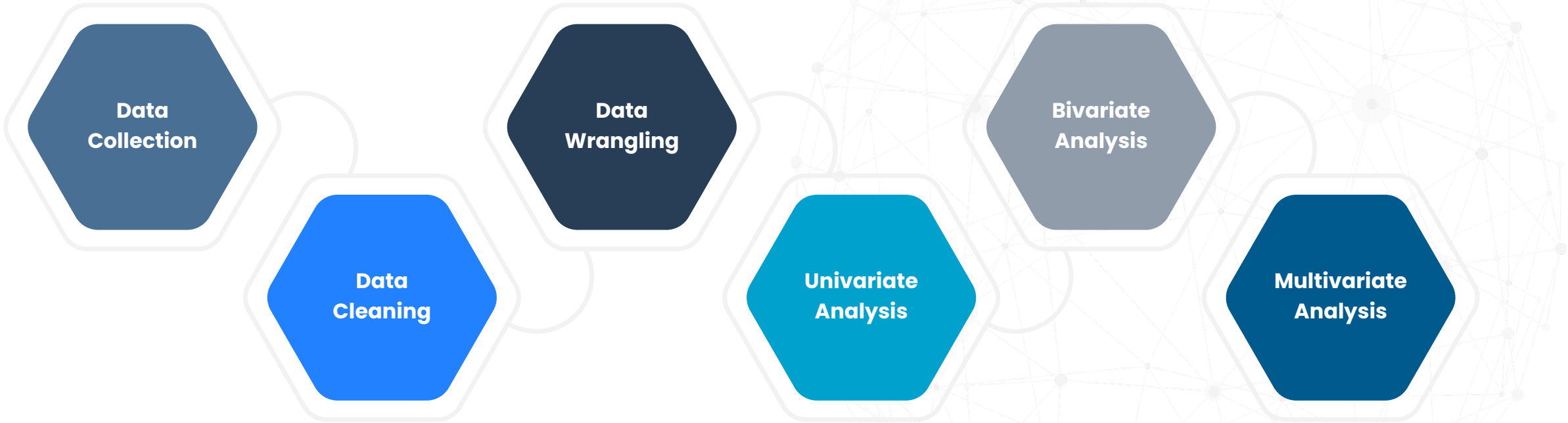
What is Exploratory Data Analysis?

- Exploratory data analysis (EDA) is used by analysts to analyze and investigate data sets and summarize their main characteristics, often employing data visualization methods.
- It helps determine how best to manipulate data sources to get the answers you need, making it easier for data scientists to discover patterns, spot anomalies, test a hypothesis, or check assumptions.



</talentlabs>

Exploratory Data Analysis - Steps



</talentlabs>

Business Value Generation

Helps understand if the results they have produced are correctly interpreted in the business context

Benefits stakeholders by confirming if the questions they're asking are right or not.

Seeing is Believing - unpredictable insights – ones that the stakeholders won't be able to figure out easily

</talentlabs>