# Course 5 end-of-course portfolio project overview: Automatidata

Learn about the Course 5 Automatidata workplace scenario!

The end-of-course project in Course 5 focuses on your ability to build regression models using Python. The end-of-course projects were designed with you in mind, offering an opportunity for you to practice and apply your data analytic skills. The materials provided here will guide you through discussions with co-workers, internal team members, and external stakeholders.
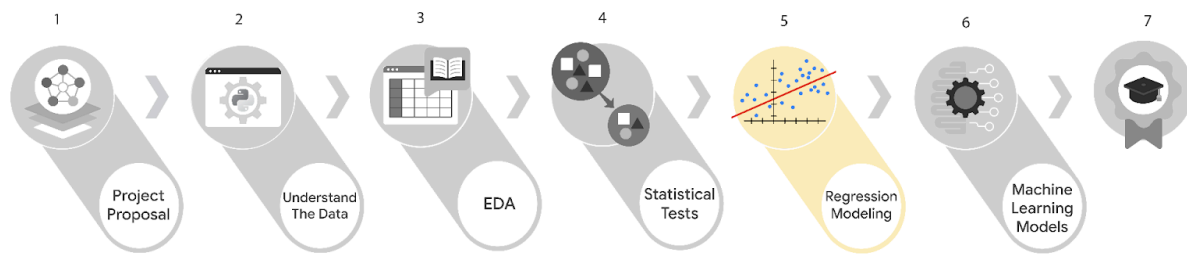
Learn more about the project, your role, and expectations in this reading.

Background on the Automatidata scenario

Automatidata works with its clients to transform their unused and stored data into useful solutions, such as performance dashboards, customer-facing tools, strategic business insights, and more. They specialize in identifying a client's business needs and utilizing their data to meet those business needs.

Automatidata is consulting for the New York City Taxi and Limousine Commission (TLC). New York City TLC is an agency responsible for licensing and regulating New York City's taxi cabs and for-hire vehicles. The agency has partnered with Automatidata to develop a regression model that helps estimate taxi fares before the ride, based on data that TLC has gathered.

The TLC data comes from over 200,000 taxi and limousine licensees, making approximately one million combined trips per day.

Note: This project's dataset was created for pedagogical purposes and may not be indicative of New York City taxi cab riders' behavior.

Project background

Automatidata is near the end of the TLC project. The following tasks are needed at this stage of the project:

- Determine the correct modeling approach
- Build a regression model
- Finish checking model assumptions
- Evaluate the model
- Interpret model results and summarize findings for stakeholders within TLC

Your assignment

You will create a regression model. Determine the type of regression model that is needed and develop one using the TLC data.

Team members of Automatidata and the New York City TLC

Automatidata Team Members

- Udo Bankole, Director of Data Analysis
- Deshawn Washington, Data Analysis Manager
- Luana Rodriquez, Senior Data Analyst
- Uli King, Senior Project Manager

Your teammates at Automatidata have technical experience with data analysis and data science. However, you should always be sure to keep summaries and messages to these team members concise and to the point.

New York City TLC Team Members

- Juliana Soto, Finance and Administration Department Head
- Titus Nelson, Operations Manager

The TLC team members are program managers who oversee operations at the organization. Their roles are not highly technical, so be sure to adjust your language and explanation accordingly.

Specific project deliverables

In this end-of-course project, you will gain valuable practice of your new skills as you complete the following deliverables:

- Complete a PACE Strategy Document to consider questions, details, and action items for each stage of the project scenario
- Answer the questions in the Jupyter notebook project file
- Build a regression model in Python
- Report the results in an executive summary

Good luck in your role! Automatidata looks forward to seeing how you communicate your creative work and approach problem-solving!

Key takeaways