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.*

Team members at 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.

Note: *The story, all names, characters, and incidents portrayed in this project are fictitious. No identification with actual persons (living or deceased) is intended or should be inferred. And, the data shared in this project has been created for pedagogical purposes.*

Project background

Automatidata is in the earliest stages of the TLC project. The following tasks are needed before the team can begin the data analysis process:

- Build a dataframe for the TLC dataset

- Examine data type of each column
- Gather descriptive statistics

Your assignment

You will build a dataframe for the TLC data. After the dataframe is complete, you will organize the data for the process of exploratory data analysis, and update the team on your progress and insights.

Specific project deliverables

With this end-of-course project, you will gain valuable practice and apply your new skills as you complete the following:
- Complete the questions in the Course 2 PACE strategy document
- Answer the questions in the Jupyter notebook project file
- Complete coding prep work on project's Jupyter notebook
- Summarize the column Dtypes
- Communicate important findings to DeShawn and Luana in the form of an executive summary

Good luck with this project! Automatidata looks forward to seeing how you communicate your creative work and approach problem-solving!

Key takeaways

The Google Advanced Data Analytics Certificate end-of-course project is designed for you to practice and apply course skills in a fictional workplace scenario. By completing each course's end-of-course project, you will have work examples that will enhance your portfolio and showcase your skills for future employers.