**Technical Test**

1. Please train a ML model in python which predicts GO Rail ridership using the attached dataset. You can have freedom to choose the model, the level of detail you want to go for and indicate your reasoning for the model/approach you choose. Do not feel limited to only use the provided data as the floor is yours what you will do with the data, and insights that you will be sharing with us.
2. Please share if you have any additional thoughts/ideas.

A formal powerpoint presentation is not required, while we’re happy to take a look if you plan to share. At a minimum you will do a codewalk through with us (please get ready to share your screen). Our goal is to assess your process (thinking, testing, pre-processing, exogenous variables you incorporated, etc); do not worry about model’s accuracy. State your assumptions as needed.

|  |
| --- |
| Terms definition:  **Ridership:** number of people boarding the train.  **Total trips:** total number of trains passing a specific station over a specific period of time.  **Time period:**  **Weekend/Holiday:** all hours of the day during which trains operate: 18 hours     AM Peak:  6:00 - 8:59     Midday:    9:00 - 14:59     PM Peak:  15:00 - 17:59     Evening:    18:00 - 23:59 |

Once your analysis is complete, please send relevant files (python code/notebook, presentation if any) to [Joshua.oh@metrolinx.com](mailto:Joshua.oh@metrolinx.com) by the morning of the interview day the latest. You’re also welcome to reach out to Joshua Oh for any questions.