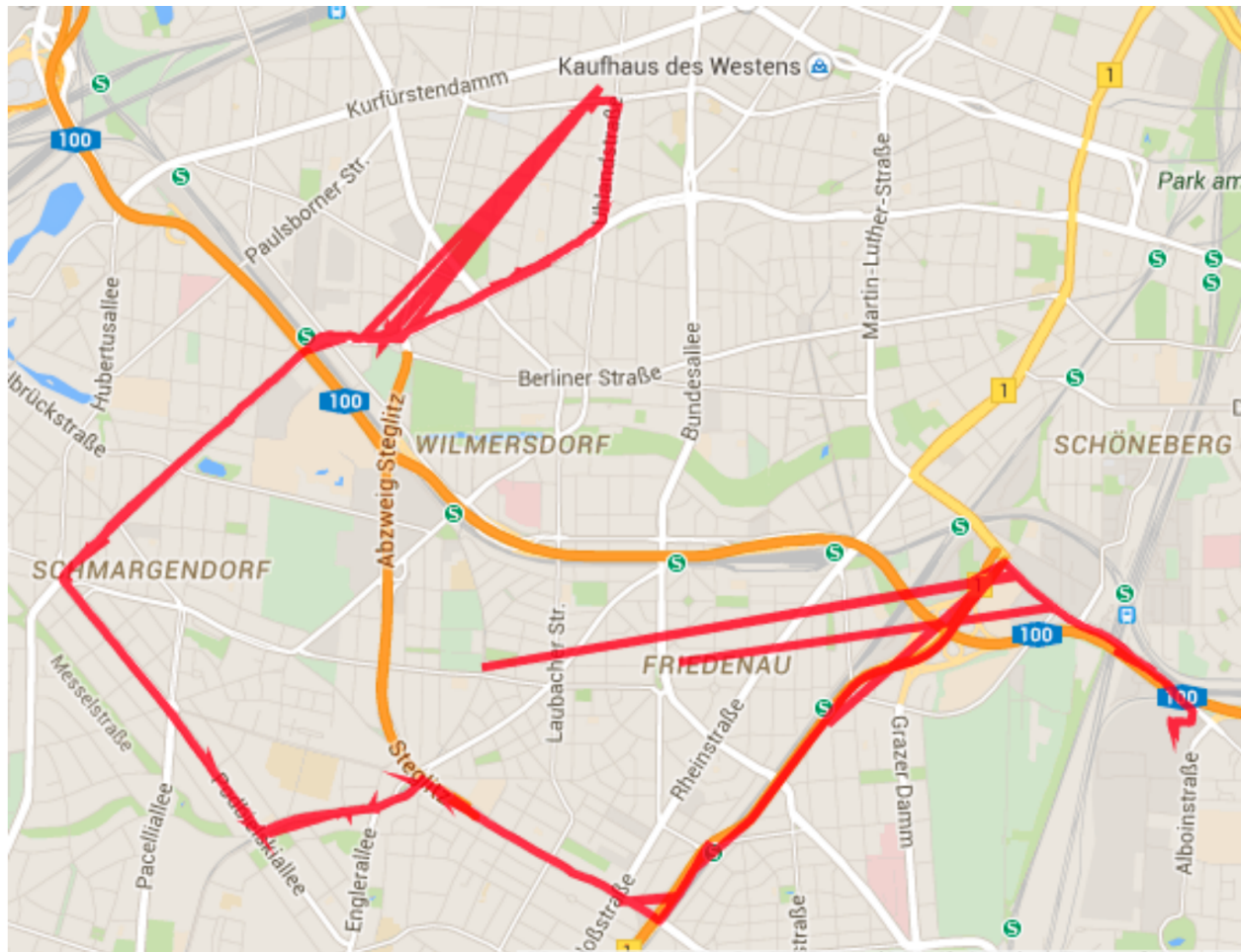# Microsoft To-Do Data Challenge

Is Data Engineering your true call? Come join the Microsoft To-Do analytics team! We've prepared two challenges for you and are looking forward seeing how you solve them. Please follow the instructions below and send your results to v-elkoiv@microsoft.com within two business days after receiving this task.

Keep in mind: there's many ways to get around the problems presented here. We are interested in your logic, so be descriptive, even verbose, and explain why you think your solution fits best!

Filter noisy GPS data

Download this dataset which describes the movement of a vehicle. Data is recorded by a sensor which naturally would contain false and noisy measurements.

1. Write a program in a scripting language of your choice that executes the following:

    • Takes the provided file as an input parameter

    • Filters out the incorrect data points

    • Outputs the original distance traveled by the vehicle including the false data points

    • Outputs the real distance traveled by the vehicle, excluding the false data points

2. Provide code and brief documentation on the logic of the algorithm

3. Describe how your solution could be further improved (if you had additional time)

Load the GeoLife GPS Trajectories dataset

Download the dataset collected by Microsoft Research Asia

1. Examine carefully the data and design a database as if you were preparing a real-life application

2. Write an ETL script in your programming language of choice that loads the data into a free-tier Microsoft Azure SQL Data Warehouse. You can sign up for a 30 day free trial here. Just like the database, your ETL solution should be production ready.

3. Provide code and documentation

Tell us about yourself (Optional)

Do you have an open source project you could share with us or anything exciting going on in your Github repo?

Show us something you've built or contributed to. It's okay if you don't have anything public: describe a project that you're proud of, something you've worked on really hard, a challenge that pushed your limits and made you a better engineer. We would love to hear about your experiences.

Send your results to v-elkoiv@microsoft.com and don't forget to have fun while tackling the challenge.