

# Data Science Work Sample

Root Insurance Co.

Feb. 2, 2018

This is the Data Science work sample for the Data Analyst position at Root Insurance Co. For all problems, complete in your language of choice (though R or Python are preferred) and submit a commit history via Github. Note for Problem 2, an install of R is required for troubleshooting, regardless of what language is used for analysis.

The units for the datasets are as follows:

- accuracy: meters (lower value = more accurate)
- heading: degrees
- latitude: degrees
- longitude: degrees
- speed: meters per second
- timestamp: epoch seconds

**Problem 1:** Analyze the enclosed trip data (`sample_trip.csv`) and provide the following summary details (showing work via commented code):

- Total trip duration [s]
- Total trip distance, computed in two ways [mi]
- Turns larger than 60 degrees (count, as well as their location and time)

*Note, be on the lookout for all sorts of bad data, including missing data, suspicious values, strange formatting, etc.*

**Problem 2:** Run the enclosed R script (`mock_function.R`) and troubleshoot the code to eliminate the errors and/or bad values. If possible, clean the data so that the code gives reasonable results. In addition, comment on the two subfunctions, whether or not they are effective in completing their specified purpose, and how they could be improved.