

Hippo Hacks 2019

Each of us have used one or more of the many transit options in Washington, DC, from the WMTA metro, the bus, uber, Capitol Bikeshare, or even the Dockless scooters. Through various public and private efforts, WMTA, DDOT, and others have worked to make DC's transit data available to the public.

Microsoft has partnered with GW DATA for their third annual Hippo Hacks data science hackathon to bring together students interested in coding, statistics, and data to work on cool data hacks! For this year's competition, you are encouraged to work in teams to leverage DC's open transit data to solve some of the biggest transportation challenges that DC residents like yourself face on a day to day basis.

This challenge is intentionally left open-ended so you have the chance to leverage this data to solve the challenge you are passionate about, get creative, and have something unique to show to our judges. Data science competitions such as these are places to practice your skills in data analysis, machine learning, data visualization, great opportunities to build up your GitHub profile, and meet others. All skillsets are welcome!

The output can be any of the following:

- Data Visualization / Dashboard
- Web Application
- Predictive Model
- PowerPoint Presentation
- Bot
- Application
- GIS / map
- And more!

You are encouraged to use any tools that you are familiar with but here are some suggestions (just pick the one or ones you're familiar with our use the hackathon to stretch and try something new!)

- Statistical Programming Languages:

- R: <https://www.rstudio.com/>
 - R Shiny for interactive R visualizations: <http://shiny.rstudio.com/>
- Python: <https://www.python.org/>
 - Flask for building a Python web app: <http://flask.pocoo.org/>
- Programming Languages
 - C#: <https://visualstudio.microsoft.com/>
 - If C# is your language of choice, ML.Net is a machine learning library built for C#: <https://dotnet.microsoft.com/apps/machinelearning-ai/ml-dotnet>
 - Java
 - JavaScript/TypeScript
- Data Visualization:
 - Power BI: <https://powerbi.microsoft.com/en-us/>
 - Tableau: <https://www.tableau.com/>
- Mapping/GIS
 - ArcGIS: <http://www.arcgis.com/index.html>
- Bot
 - Azure Bot Service
 - You can sign up \$200 free azure account (no credit card required)
- And more!

Data Sources

- WMTA Bus & Metro Data: <https://developer.wmata.com/docs/services>
 - Data on metro and bus routes/schedules, incidents, and bus/train predictions
- DC Circulator: <http://www.dccirculator.com/circulator-data/>
 - Performance and usage data for the DC Circulator bus
- DC Streetcar: <https://www.dcstreetcar.com/about/resources/developer-resources/>
 - Schedules, routes, and fares for the DC streetcar
- Dockless Rideshare API: <https://ddot.dc.gov/page/dockless-api>
 - Dockless bikeshare and scooter share locations, battery levels, vehicle types
- Uber DC Transit Times: https://movement.uber.com/explore/washington_DC/travel-times/
 - Uber ride average travel times
- OpenData DC: <http://opendata.dc.gov/>
 - Repository for DC open data from demographics to public safety to transportation and more
- Feel free to choose any from these or combine with outside data sets!

The next page contains a copy of the judging criteria for your reference.

Judging Criteria

A panel of judges will evaluate each project and pick winners. The judges will base their assessments on the 3 main categories below and award prizes to the top project in each category.

Product

A data product, something used by a customer over time, will be judged upon its usefulness, interface and creativity.

1

2

3

4

5

Analytics

A good analysis project will apply sophisticated statistics or machine learning to derive interesting insights from the dataset. An analysis project will be judged upon technical strength, complexity, impact of derived result and innovation.

1

2

3

4

5

Visualization

A good visualization project will find interesting ways to consider a dataset or combine information from multiple data sources. A visualization project will be judged upon design aesthetics, value add, and complexity.

1

2

3

4

5

Overall Score: _____

Feedback: