Open Data
User Journey Analysis

Samuel Engelberg - Mayor’s Office of Data Analytics
April 2018
• Overview
• Key Findings
• User Breakdown
  1) Explorers
  2) Drivers
  3) Test Takers
  4) Fact Checkers
• Key Takeaways
• Next Steps
We're looking to answer the following questions:

- Who are Open Data users?
- What sources are referring them to our site?
- What is the first page they look at?
- And where do they go next?
Previous Research and Findings

The Open Data Team engaged Reboot in Spring 2017 to research the users and opportunities for engagement with them.

- Captured information through 20 different interviews across various different boroughs and industries – categorizing them into different user profiles
- Great qualitative data, we were curious to know what the quantitative data is as well
- You can find the research done by Reboot on the Open Data website under “About” and “Laws and Reports” in the “Open Data User Research” section

This analysis builds off of Reboot’s work by exploring insights around user behavior found in website traffic to the Open Data website.
Terminology

**Google Analytics** is a web analytics service that tracks and reports (anonymized) website traffic.

**Session** is a group of user interactions with a website that take place in a given timeframe.
Key Findings

1. 2 million Sessions (March ‘17 - March ‘18)
2. 35% Mobile
   62% Desktop
   3% Tablet
3. Top referral sources
   1) Google
   2) Nyc.gov
   3) Direct
4. 75% drop-off/exit after viewing 1 page
Understanding our Users
After digging into the user data we were able to identify and group users into the following groups, based on the Assets they viewed.

The “other” user group consists of many smaller groups.

- Based on 2 million Sessions from March ’17 - March’18
- *Google Analytics Behaviour Flow Report
Fact Checkers: 16% of sessions

NYC residents visiting the site, directly digging into the data

- Coming directly to a page finding what they need and then leaving it.
- Fact Checkers are: conducting research, identifying trends in their area, or working on a specific case/project
  - Looking at highly unique data sets with no particular use cases
Fact Checkers: 16% of sessions

Some of the most popular Assets that Fact Checkers look into are the following:

- NYPD Motor Vehicle Collisions
- 311 Service Requests from 2010 to present
- Film Permits
- Subway Entrances
- OATH Hearings Division Case Status
- Citywide Payroll Data Fiscal Year
- Average Salaries In Department of Correction
Drivers: 37% of sessions

02

Checking assets that are provided by the Taxi and Limousine Commission (TLC)

- The TLC is responsible for licensing and regulating NYC’s medallion (yellow) cabs, for-hire vehicles, commuter vans and paratransit vehicles
- The TLC has been publishing on NYC Open Data before Local Law 11 of 2012 was in place
Use Cases

- Drivers refer to open data to check up on:
  - Licenses, Vehicles, and Applications
  - Many drivers have shortcuts to datasets e.g. bookmarks
  - External stakeholders reference data for Driver authentication
Leverage NYC Open Data to check civil service exam results

- Unions & Supervisors at other Agencies specifically check: civil service test results
- This data can be found in assets such as:
  - Civil List, Civil Service Titles, City Record Online, and Civil Service List Active
- Department of Citywide Administrative Services (DCAS)
  - Responsible for publishing this data and managing exams
  - Advertises datasets on City Record Online and Green Book Online
Explorers

04

Exploring the site content and data catalog

- Look into Home/ Info pages such as: home, login, how to, management report and dashboard
- Browse and search the Data Catalog
  - Delve into a diverse set of filters, searches and over 60% of users continue to other pages
  - Most popular browsing queries *see next slide
Data Catalog Browsing

Explorers most commonly used the following search queries & filter options while browsing the Data Catalog.
The most common referrals for all users are: direct, Google and nyc.gov.

Drivers and Fact Checkers have very few Users that navigate the site after viewing one page, while Explorers and Test takers are more likely to look at other assets and pages.
Key Takeaways

With Drivers, Fact Checkers, Explorers and Test Takers making up the majority of users the following considerations need to be made to enhance strategy and engagement.

01 | Offer additional resources for Drivers (e.g. maintenance status)
02 | Recommend similar datasets to Fact Checkers
03 | Ensure usability of exploring by emphasizing most common queries
04 | Determine what other assets Test Takers are viewing
Future opportunities for user research

01 User Interviews & Surveys
Identify individual users and their use cases of Open Data and their unique journey on the platform. Through on Website surveys or invitations to interviews.

02 Google Analytics
Automate research to allow up to date findings and changes in behaviour. And drill down on pages with high traffic as well as long time spend.

03 Agency Feedback
Collaborate with agencies to better understand their users and potential use cases. Specifically referral sources and Agency commitment to Open Data,
Thank you

For any questions related to this report contact the Open Data Team via the Contact Us page on [www.nyc.gov/opendata](http://www.nyc.gov/opendata)