HEAVY.AI

HeavyCloud Trial Quick Start Guide

v1.0

Updated: September 2024

Notices

HEAVY.AI, Inc. (HEAVY.AI) may not offer the products, services, or features discussed in this document in all countries or to any particular users. HEAVY.AI may have patents or pending patent applications covering subject matter described in this document. The furnishing of this document does not grant you any license to these patents or to any of HEAVY.AI's intellectual property.

HEAVY.AI PROVIDES THIS DOCUMENT "AS IS" WITHOUT WARRANTY OF ANY KIND, EITHER EXPRESS OR IMPLIED, INCLUDING, BUT NOT LIMITED TO, THE IMPLIED WARRANTIES OF NON-INFRINGEMENT, MERCHANTABILITY OR FITNESS FOR A PARTICULAR PURPOSE. Some jurisdictions do not allow dis- claimer of express or implied warranties in certain transactions, therefore, this statement may not apply to you.

Any references in this document to non-HEAVY.AI products, services, or websites are provided for convenience only and do not in any manner serve as an endorsement. Information concerning non-HEAVY.AI products was obtained from the suppliers of those products, their published announcements or other publicly available sources. HEAVY.AI cannot confirm the accuracy of performance, compatibility or any other claims related to non-HEAVY.AI products.

Statements regarding HEAVY.AI's future direction or intent are subject to change or withdrawal without notice, and represent goals and objectives only.

HEAVY.AI, the HEAVY.AI logo, and heavy.ai (website URL) are trademarks or registered trademarks of HEAVY.AI, registered in the United States and other jurisdictions. Other product and service names might be trademarks or service marks of HEAVY.AI or other companies.

Version	Date	Notes
1.0	1 September 2024	

Table of Contents

Introduction

If you are reading this document, you have successfully signed up for the FREE HeavyCloud Trial and have already logged into HEAVY.AI, or you are getting ready to login.

This document is intended to provide you with more details on what data and dashboards have been pre-loaded in this environment.

Our recommendation is to spend some time understanding the data and dashboards that have been provided and then consider loading additional data. You can load up to 100 million rows of additional data along with 3 total users. Don't hesitate to reach out to support@heavy.ai if you need help along the way.

This HeavyCloud Trial environment that you will be using is a small trial system. We are trying to provide you with a taste of the HEAVY.AI platform. The data provided is the type of information you would use if you were doing a site suitability analysis (think of opening a new retail store location). We understand that the data we have provided is a fraction of what you would use if you were truly executing a proper site selection process. Our customers load BILLIONS of records and experience lightning fast query response times.

For purposes of this quick start guide, we are providing some basic information. This video provides an overview as well. Additionally, we have provided a summary within the demo itself. If at any time you are looking for more detailed information and/or you have questions, don't hesitate to utilize the HEAVY.AI Community, Support or send an email to support@heavy.ai.

Upon logging in, open the dashboard shown in the list below. The buttons on the left can be used to explore the different aspects of the platform.
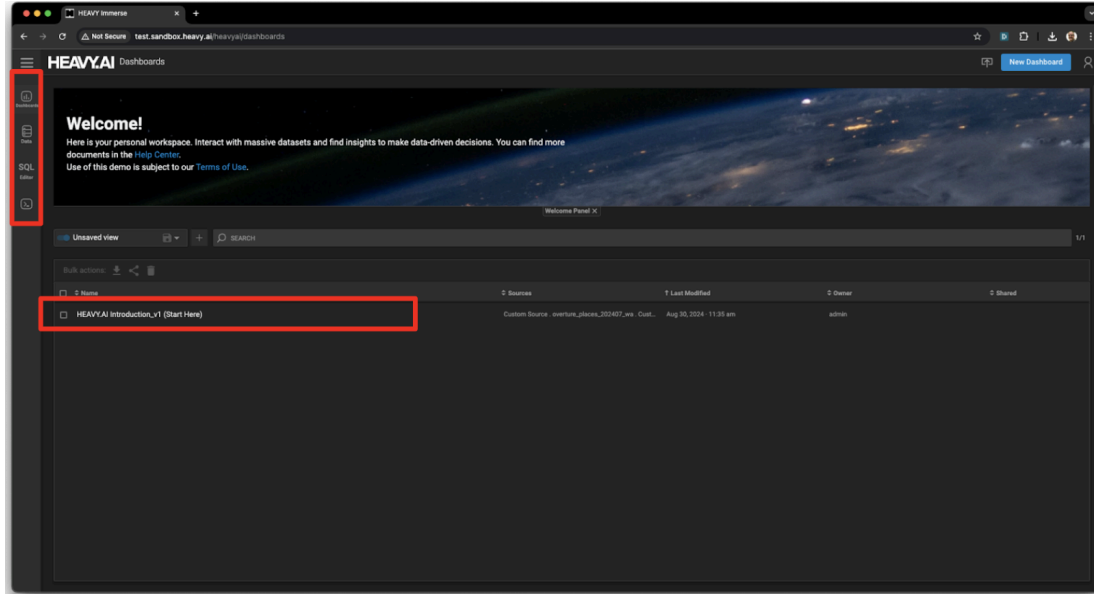
From within the dashboard, use the tabs on the lower left to navigate to the various pages.
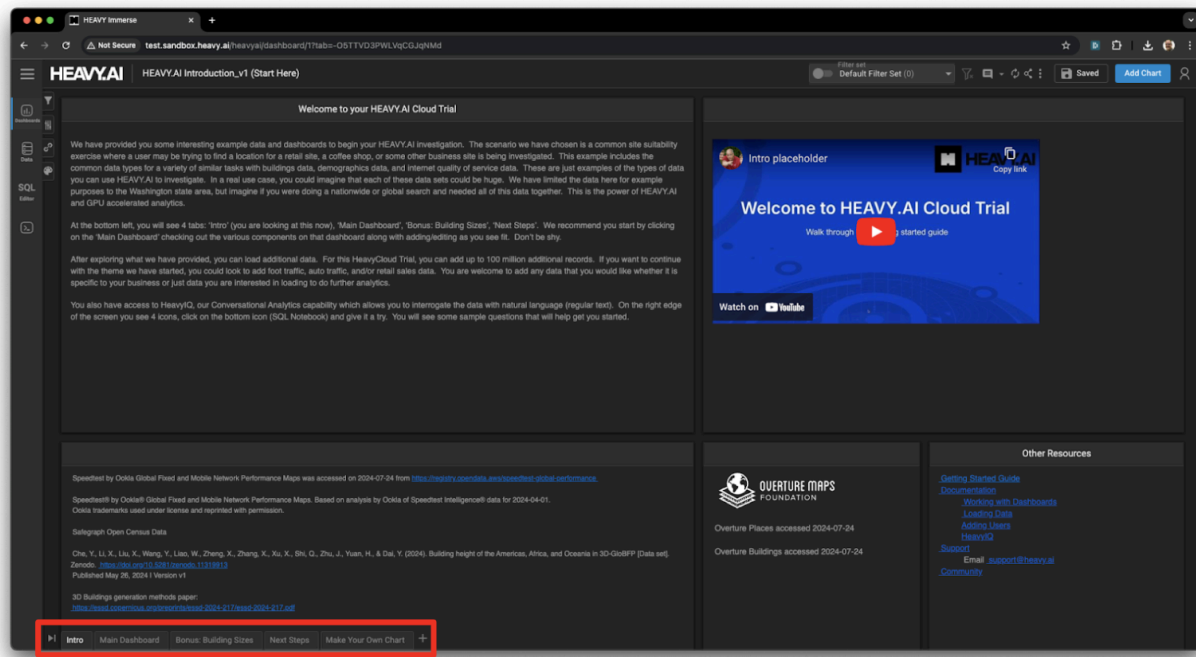
This interface, HeavyImmerse, provides dynamic interactive visualizations. Within the HeavyImmerse dashboards, you can instantly cross-filter billions of observations with no downsampling required. Additionally, you can visualize huge geospatial datasets interactively with HEAVY.AI's native, GPU-accelerated rendering engine.

Data

Data has been preloaded from the following original sources:
- [Census.gov](#) - data utilized is from [SafeGraph](#)
- [Overture Maps Foundation](#)
  - [Buildings](#)
  - [Places](#)
- [OOKLA](#)
  - [Fixed Wireless Statistics](#)

There are approximately 5.7 million rows of data already loaded in your environment, which by HEAVY.AI standards is very small. The preloaded data is limited to the US state of Washington. You also have the ability to load up to 100 million additional rows into this environment.

Raw data was accessed from the data sources listed above. Then, some joining and manipulation was done in order to be able to look at the data/layers on maps and different dashboard components. As with most data, decisions were made to cleanse the data along the way. Additionally, you may notice some data quality issues. The intention is to provide you with some exposure to the HEAVY.AI platform with an interesting sample dataset that most people will be able to relate to.

The simplest way to think about the data is that 'Places' can have a many:1 relationship with 'Buildings' (for example, a Shopping Mall). Buildings exist within a Census Block, which exists within a Census Block Group, which exists within a Census Tract. Associated with the Census Block Group are various attributes like population and average household size. Tracts additionally include more sensitive data like median household income.

Crowd sourced data like Ookla's open data has also had some privacy protection applied. In their case, they release raw metrics for about a 5 by 5 block area. These metrics were interpolated down to building resolution, and a weighted index was built combining upload, download and latency metrics (60%-30%-10%). This is a common technique for dimensionality reduction, but of course appropriate weights might vary for some audiences. Video streamers, for example, do care about upload speed more than the average person, and gamers care about latency.

The location of current coffee shops is being shown (and other selected types of businesses) along with applicable household income and internet access quality across the state of Washington to assess potential locations for new coffee shops. Understanding this is an incredibly basic site selection example, you can choose to add additional layers (like traffic - foot, vehicular) to get additional perspective on the power of this analytic platform which is being used by customers across a wide range of industries including Telco, Utilities, Oil & Gas, and Public Sector, to name a few.

Main Dashboard
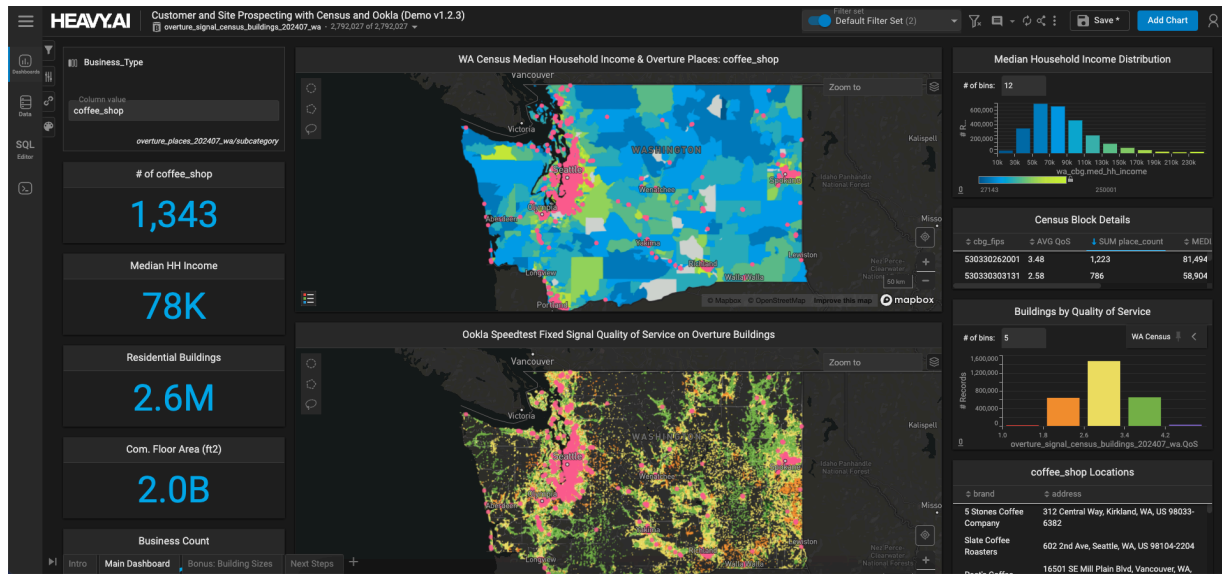
From most of the components on the Dashboard, you can click on the ⚙ in the top right corner of any given component to be able to experiment editing.

Along the left edge of the dashboard is a collection of Number Charts. The maps in the center both have a point layer (Places) on top of a polygon layer (Census Block Groups on the top map, Building Footprints on the bottom map). Along the right edge of the dashboard is a series of different charts including Combo Charts and Tables.

Dashboard Components:

- Business Type

    - Click in the 'Column value' field to be able to select different Business Types.

- # of <Business_Type>

    - Displays the number of the selected business type which are currently displayed on the maps.

- Median HH Income

    - Displays the median household income for the area currently shown on the maps.

- Residential Buildings

    - Displays the number of residential buildings currently shown on the maps.

- Com. Floor Area (ft2)

- ○ Displays the square footage of all the commercial buildings currently shown on the maps
- Business Count

  ○ Displays the number of businesses currently shown on the maps.

- WA Census Median Household Income & Overture Places: <Business_Type>

  ○ Displays the median household income by Census Block Group (CBG) and the applicable business based upon the business type selection.
  ○ You can scroll in/out of the map which will automatically update other dashboard components.

- Ookla Speedtest Fixed Signal QoS on Overture Buildings

  ○ Displays the quality of service for fixed wireless for each building on the map. Click on the legend in the lower left corner to see more details.
  ○ You can scroll in/out on the map which will automatically update the other dashboard components.
  ○ We have only included Fixed Wireless measurements in this demo. Fixed wireless is a type of wireless broadband that uses radio waves (as opposed to cabling) to connect to the internet from a home or business.

- Median Household Income Distribution

  ○ This shows the household income distribution for the CBGs shown on the maps.
  ○ CBGs will show if the applicable centroid is on the visual portion of the displayed map.
  ○ You can highlight one or more of the bars to focus the maps on CBGs with the associated income range.

- Census Block Details

  ○ Provides overall data for the CBGs shown on the maps.
  ○ You can select specific CBGs to see where they are located on the map.

- Buildings by Quality of Service

  ○ Shows the quality of service distribution for the buildings shown on the map.
  ○ You can highlight one or more of the bars to focus the maps on buildings with the associated quality of service.

- <Business_Type> Locations

  ○ Provides details for the applicable businesses shown on the map based upon the <Business_Type> selection.

Enjoy investigating and working with the various components on the dashboard. Feel free to add, edit, and delete components and dashboards as you see fit. Here is some additional documentation that describes how to create/edit dashboards.

Another dashboard that you can explore is called "Bonus: Building Sizes." This dashboard shows the height of buildings as modeled using a machine learning algorithm. Building heights, of course, are the key to estimating a key concern in both real estate and suitability modeling - floor area. In addition to authoritative and crowd-sourced data, this represents a third type of data source increasingly widely available. It is incredibly comprehensive, estimating the height of literally every building in the world. However this inevitably means error checking is difficult, and there are indeed some errors apparent if you look closely. While HEAVY.AI will not magically make your data perfect, it will allow you to include and inspect a wide variety of datasets, many of which like this one are quite huge. You can also selectively merge large data, which we've done in the main dashboard by estimating heights and then floor areas from this source only when Overture Foundation did not provide any height.

HeavyIQ

HeavyIQ provides conversational analytics on your massive data. It takes advantage of the latest LLM (large language model) technology so that you can ask questions of your data in natural language and get back actionable visualizations.

From within the HeavyImmerse interface, you can see the tables and columns as well as some sense for the data within each column via the SQL Notebook.
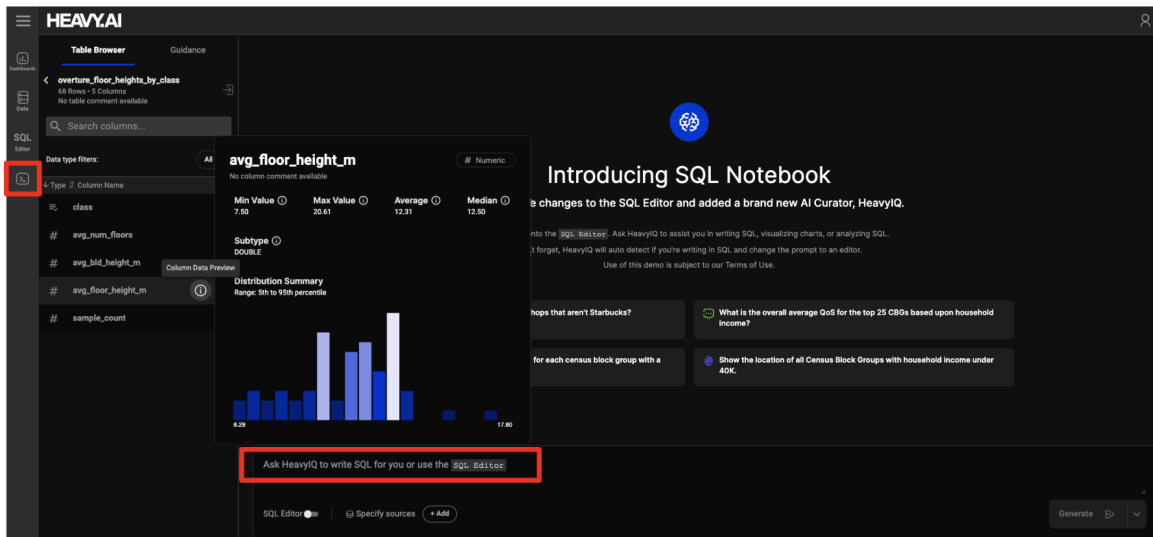
From within the SQL Notebook, you can submit prompts/questions via natural language (regular text) to get answers. An example, as you are getting to know the data, may be . . . "What

values exist for class on the overture buildings ookla table that don't exist in the overture buildings places table?" BTW, this is taking advantage of a large language model to do the text-to-SQL translation behind the scenes.

Here are some sample prompts/questions to get you started exploring Conversational Analytics:

- What is the average QoS of all coffee shops that aren't Starbucks?

- What is the overall average QoS for the top 25 CBGs based upon household income?

- What is the name and address of the building with the most floors? Also include the number of floors.

- What is the average household income for each census block group with a Dutch Bros Coffee?

- Where is CBG 530350802002?

- Show the location of all Census Block Groups with household income under 40K.

When you see HeavyIQ answering your questions incorrectly, try adding snippets via Guidance to help the LLM better understand how to answer your questions. You will see some snippets already loaded in this application.
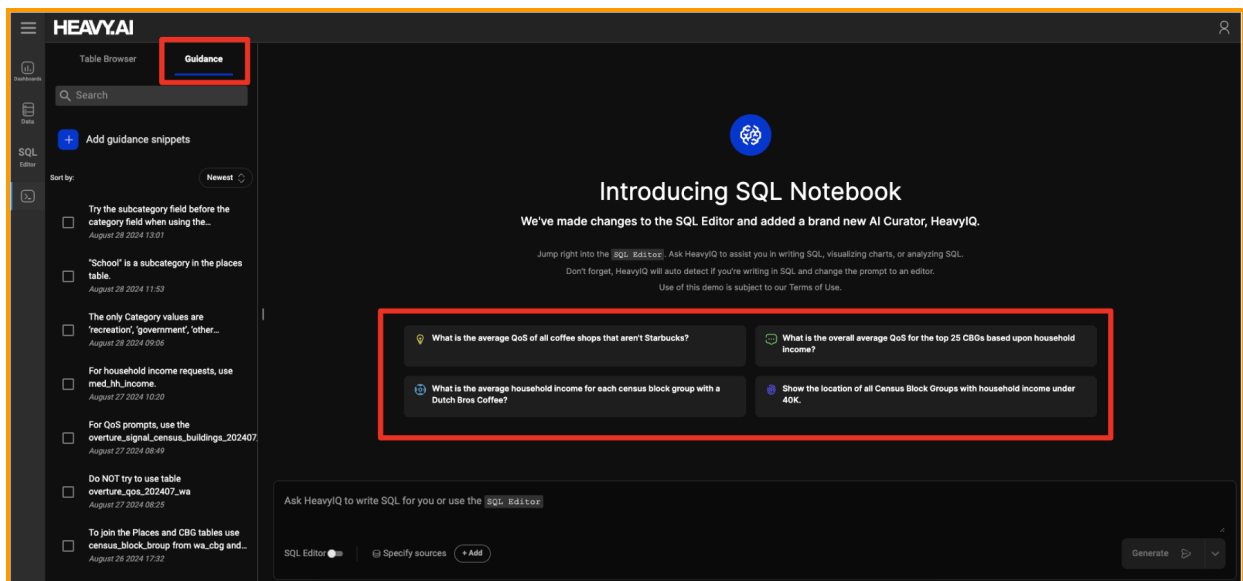
Next Steps

We encourage you to load your own data to build additional dashboards and explore more with HeavyIQ. Click here to learn about options for loading data.

Need Help?

If you have any questions while using this Quick Start Guide, you can submit a question through our Community Forums or a request through our Support Portal.