Student Onboarding Analysis Project Write-Up

Author: Avinash Bisram

Sections:

- Overview
- Background
- Dataset
- Goals
- Deliverables
- Key Performance Questions
- Additional Insights and Recommendations

Overview: Hi! The purpose of this write-up is to explain the background, goals, deliverables, and insights gained from the *Student Onboarding Analysis in Tableau* project. This write-up is meant to accompany the final Tableau Dashboard which **can be found here** (a screenshot of the final view can also be seen in the *Deliverables* section).

Background: This project is one of many that can be found on **365DataScience.com**. The data used was a subset of real information collected from a questionnaire during the student onboarding process (see the *Dataset* section for more information). This survey asked various questions relating to basic identifying information, student goals, how users discovered the platform, and what they hope to learn on the site. This data is invaluable to a company as they can use this information to better cater content to their most prominent demographics, understand the relative success of new user acquisition across different advertising streams, and continue to shape the platform in the future to ensure success and longevity.

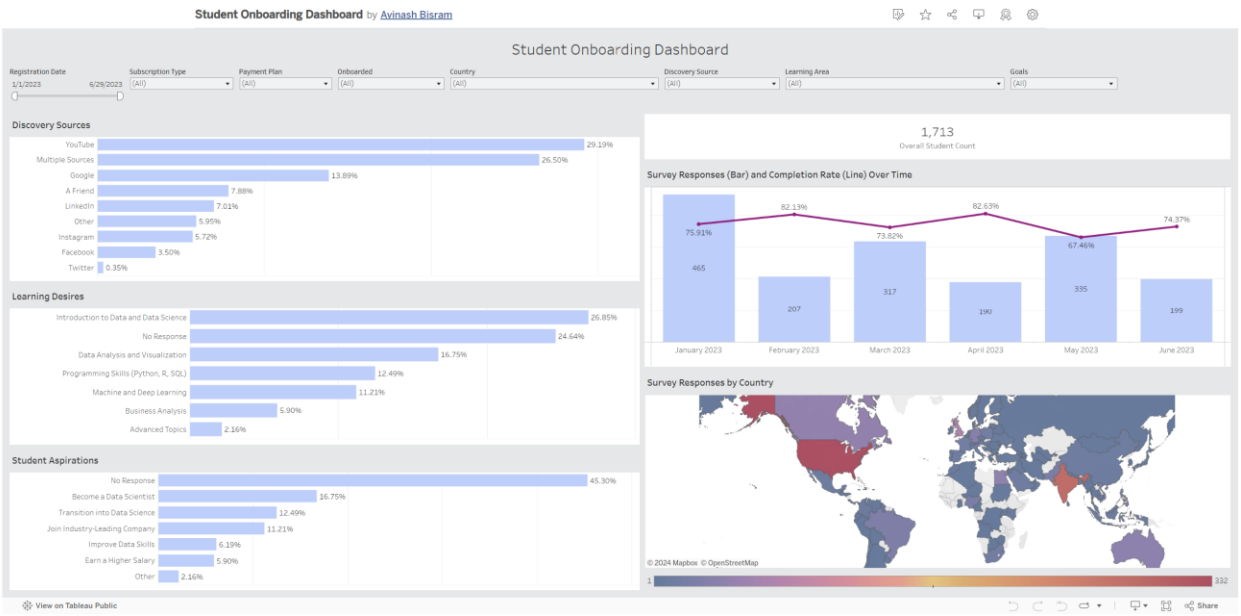
Dataset: The dataset used to construct the final dashboard was provided by 365DataScience in the form of a CSV file. This form contained preprocessed and anonymized data on real user responses to an onboarding questionnaire in the first half of 2023. The size of the data was 1713 rows and 10 features relating to various user attributes and responses including whether a user registered for a free or paid subscription, the payment renewal frequency, each user's country, as well as a few answers relating to how they heard of the platform, what they hope to learn, and the overarching goal they are trying to achieve by learning on the platform. This CSV was used as the sole data source to construct the Tableau dashboard.

Goals: The goal of this project was to create visualizations capable of answering predefined Key Performance Questions (KPQs) outlined by 365DataScience and combine them into a Tableau dashboard with various filters to uncover additional insights and make recommendations on how to

further improve the platform based on the different trends that may be uncovered. There were no explicit measures that were defined to be optimized or improved so the recommendations could take many different routes depending on which areas of opportunity were identified through these visualizations (a few that I personally found interesting are explored in more detail in the *Additional Insights and Recommendations* section).

Deliverables:

The main deliverable for this project was a Tableau dashboard created using the data provided. An image of the final view can be seen below.



The dashboard consists of **5 visualizations** (and 1 additional displayed KPI) as well as **8 filters**. **3 calculated fields** were also made to facilitate creating these visuals.

Short descriptions of the dashboard components are as follows:

- "Discovery Sources" (Top Left): Horizontal bar chart displaying the relative percentages and overall distribution of where each user heard about the platform (YouTube, Google, Facebook, etc.). Hovering over the bars display the total number of responses for each category as well.
- "Learning Desires" (Middle Left): Horizontal bar chart displaying the distribution of responses to a question asking about their main learning goals. 365DataScience.com offers numerous courses and resources that cover a breadth of skills outside of just data science.
- "Student Aspirations" (Bottom Left): Similar to the two other horizontal bar charts, this visual captures the distribution of responses for user goals.
- "Overall Student Count" (Top Right): This is a simple KPI that displays the number of responses being aggregated by the dashboard at any time (it updates as viewers interact with the different filters above it).

- "Survey Responses and Completion Rate over Time" (Middle Right): Dual-axis chart consisting of a bar chart which displays survey responses or user registrations over time and a line chart indicating the "completion rate" (how many students complete the entire onboarding survey after registering). The default view for this chart shows these metrics by month but can be drilled up or down as well (to view yearly, quarterly, etc.).
- "Survey Responses by Country" (Bottom Right): Colored map which reflects the number of registered users (or survey responses as these seemed to have been used interchangeably by 365DataScience) for each country.
- Filters: Several filters were added to enable additional analysis and levels of granularity for the dashboard including filtering by registration date, the type of subscription (free or paid), payment plan, country, goals, learning desires, and acquisition channels.

Key Performance Questions:

As the dashboard suggests, each visualization relates to a different key performance question outlined by the CEO. These were general guiding questions such as "What is the monthly survey response rate and completion rate among students?" and "What are students' desired learning areas? What is the primary or least desired learning area?". In addition to these, several questions were asked to prove the successful completion of the dashboard:

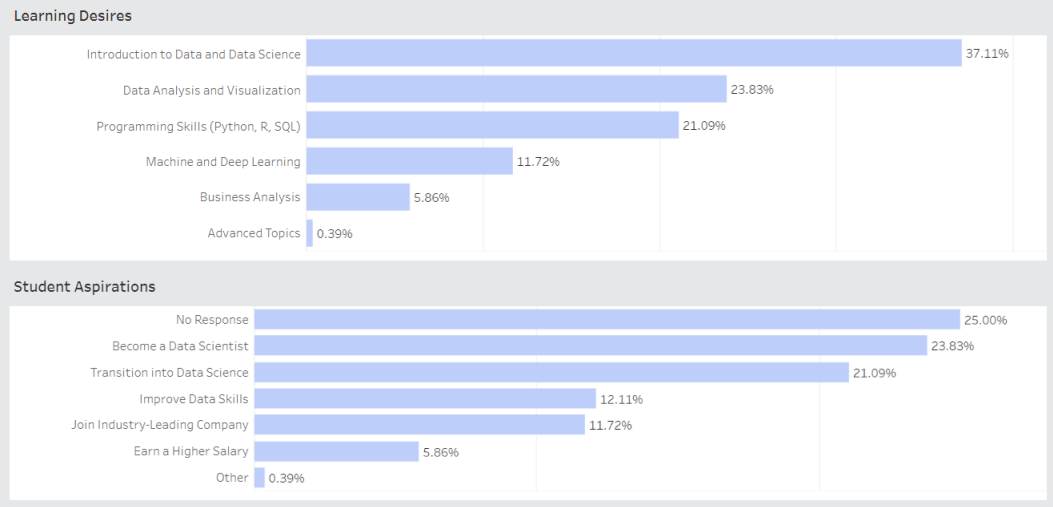
- What is the completion rate of the survey—i.e., are the survey results a representative number of the population?
    - By drilling up on the dual-axis chart to view all of 2023, we see that about 75% of all registered users were successfully onboarded (completed the entire survey). We can say that this is a "representative number" of the population as it is the majority however it would be best to define an ideal threshold before constructing the dashboard in a real business case.
- Where are the survey participants from? Which countries have the most participants?
    - As can be seen in the map chart, there were survey participants from around the world such as North and South America, Europe, Asia, parts of Africa, and even Australia (sadly there no responders from Antarctica). The two countries with the most participants can be seen colored in red as the United States of America (about 19% of all participants) and India (just under 17%). Together, these two countries represent over a third of registered users in 2023 (based on the sample of data provided).
- What is the most effective acquisition channel?
    - The most effective channel was YouTube accounting for 29% of all responses (as can be seen in the top left chart).
- What is the most common client's profile: beginner, advanced learner, etc.?
    - This is a bit harder to quantify without additional information but we can assume that "beginners" are those that stated their learning desires are being introduced to data and data science or learning programming skills such as Python and SQL. Those two categories combined make a little under 40% of the total responses so

we can say (without additional information) that the most common client are beginners.

Additional Insights and Recommendations:

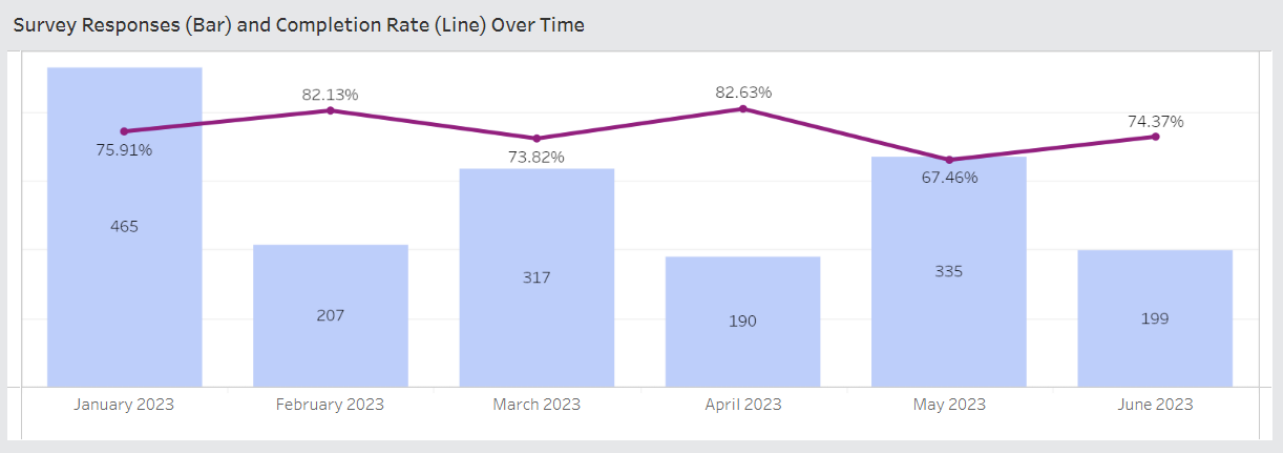
As mentioned previously, this dashboard was not constructed with an explicit goal such as how to increase the completion rate over time, creating more popular courses, etc. so the analysis and insights can take many different routes.

One direction can relate to capitalizing on the platform's most favorable demographics. If we filter the dashboard to only display results for the United States and India, we see that these two groups combined had a 99% completion rate in 2023. This number remained relatively stable over each month. Over 40% of these users claim to have come from YouTube so we can filter the dashboard further and see what their main learning desires are.



As can be seen in the image above, the top three learning desires are being introduced to data, learning data analysis and visualization, and learning foundational programming skills. Significantly less are interested in Machine and Deep Learning or Advanced Topics so it would be beneficial to ensure courses are being developed (or updated) to cover the really important information in those categories. Furthermore, a little over 40% are trying to become data scientists or transition into the field so another avenue for content creation can include interview preparation, building a professional network, and steps to land a data science job. Perhaps the content creation team on YouTube can even explore dubbing their most popular videos in Hindi to cater to the Indian audience who may be interested in these concepts but are not proficient in English. We would expect that the more people we reach and offer useful content to on YouTube, the more people would eventually subscribe to the real platform to continue their learning.

Another direction that can be taken in this analysis is strategizing how to acquire more customers overall.



As can be seen in the chart above, there is pretty clear seasonality with more users being acquired on odd months (January, March, May) and higher completion rates in subsequent months. Perhaps the platform might want to look into running sales or offering certain courses at a discount in some of these low months to make customer acquisition more stable over time.

Lastly, the platform may be interested in improving the user experience for free users in the hopes that more of them eventually convert to paid subscriptions. By filtering the dashboard for the "Free" subscription type, we see that over 40% of all registered users do not respond to the Learning Desire or Student Aspirations/Goals question. This is quite an alarming statistic and may prompt the platform to try to understand why so many of them do not respond. Perhaps they can incentivize free users to answer those questions by giving them temporary access to an otherwise "premium" course depending on their responses to those questions. For example, if a free user says they are most interested in being Introduced to Data Science, the platform can offer them unrestricted access to the **Introduction to Data and Data Science** course or offer an exclusive discount on a bundle that contains that course for a limited time. This would be a great incentive for free users to experience more of the platform (and eventually convert to a paid user) while providing the platform with invaluable data to improve the user experience for those users. A similar thought process can be followed for paid users belonging to each of the payment types to increase the probability of upgrading their memberships by understanding what they value the most.