

# Suraj Udasi

## Business Analysis with Structured Data

The screenshot shows the MySQL Workbench interface with the 'Administration - Server Status' window open for the instance 'su\_RDS'. The window is divided into several sections: 'Connection Name', 'Host', 'Socket', 'Port', 'Version', 'Compiled For', 'Configuration File', 'Running Since', 'Available Server Features', and 'Server Directories'. The 'Server Status' is 'Running'. The 'Host' is 'ip-10-1-4-205'. The 'Socket' is '/tmp/mysqlsock'. The 'Port' is '3306'. The 'Version' is '8.0.33 (Source distribution)'. The 'Compiled For' is 'Linux (x86\_64)'. The 'Configuration File' is 'unknown'. The 'Running Since' is 'Sat Nov 18 21:45:58 2023 (12 days 21:50)'. The 'Available Server Features' section shows 'Performance Schemas' as 'Off', 'Thread Pool' as 'n/a', 'Memcached Plugin' as 'n/a', 'Semisync Replication Plugin' as 'n/a', and 'SSL Availability' as 'On'. The 'Server Directories' section shows 'Base Directory' as '/rdsdbsbin/mysql-8.0.33.R4/', 'Data Directory' as '/rdsdbsdata/db/', 'Disk Space in Data Dir' as 'unable to retrieve', 'InnoDB Data Directory' as '/rdsdbsdata/db/innodb', 'Plugins Directory' as '/rdsdbsbin/mysql-8.0.33.R4/lib/plugin/', 'Tmp Directory' as '/rdsdbsdata/tmp', and 'Error Log' as 'On /rdsdbsdata/log/error/mysql-error.log'. On the right side, there are several performance metrics: 'Server Status' (Running), 'CPU Load' (---), 'Connections' (5), 'Traffic' (29.71 KB/s), 'Key Efficiency' (0.0%), 'Selects per Second' (0), 'InnoDB Buffer Usage' (14.3%), 'InnoDB Reads per Second' (0), and 'InnoDB Writes per Second' (1).

MySQL Workbench

Local instance MySQL80 x su\_RDS x

File Edit View Query Database Server Tools Scripting Help

Navigator SQL File 5 Administration - Server Status x

**MANAGEMENT**

- Server Status
- Client Connections
- Users and Privileges
- Status and System Variables
- Data Export
- Data Import/Restore

**INSTANCE**

- Startup / Shutdown
- Server Logs
- Options File

**PERFORMANCE**

- Dashboard
- Performance Reports
- Performance Schema Setup

Administration Schemas

Information

No object selected

Connection Name: su\_RDS

Host: ip-10-1-4-205

Socket: /tmp/mysqlsock

Port: 3306

Version: 8.0.33 (Source distribution)

Compiled For: Linux (x86\_64)

Configuration File: unknown

Running Since: Sat Nov 18 21:45:58 2023 (12 days 21:50)

Refresh

**Available Server Features**

- Performance Schemas: Off
- Thread Pool: n/a
- Memcached Plugin: n/a
- Semisync Replication Plugin: n/a
- SSL Availability: On

**Server Directories**

- Base Directory: /rdsdbsbin/mysql-8.0.33.R4/
- Data Directory: /rdsdbsdata/db/
- Disk Space in Data Dir: unable to retrieve
- InnoDB Data Directory: /rdsdbsdata/db/innodb
- Plugins Directory: /rdsdbsbin/mysql-8.0.33.R4/lib/plugin/
- Tmp Directory: /rdsdbsdata/tmp
- Error Log: On /rdsdbsdata/log/error/mysql-error.log

Server Status: Running

CPU Load: ---

Connections: 5

Traffic: 29.71 KB/s

Key Efficiency: 0.0%

Selects per Second: 0

InnoDB Buffer Usage: 14.3%

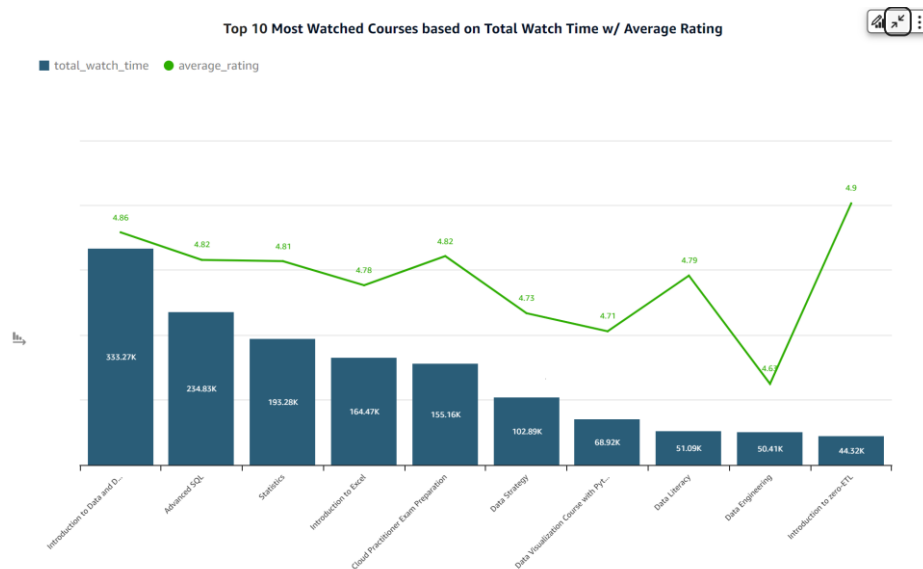
InnoDB Reads per Second: 0

InnoDB Writes per Second: 1

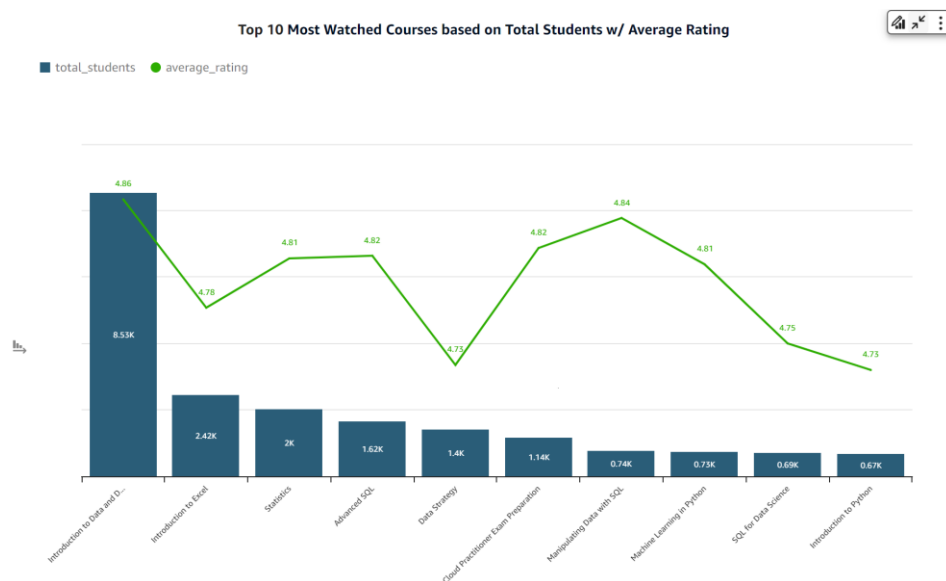
### Q1: Which courses are the most watched by students, and how are they rated?

The answer to this question is twofold, most watched could be categorized by **a. total minutes watched of the course** and **b. total number of students learning the course**. (Refer Query – A.1)

Part a. The visual below displays the Top 10 most-watched courses by total watch time, with “Introduction to Data and Data Science” topping the charts with a total of 333.2K minutes watched having an average rating of 4.86, followed by “Advanced SQL” with 234.8K minutes watched and an average rating of 4.82 and “Statistics” with 193.3K minutes watched and an average rating of 4.81.



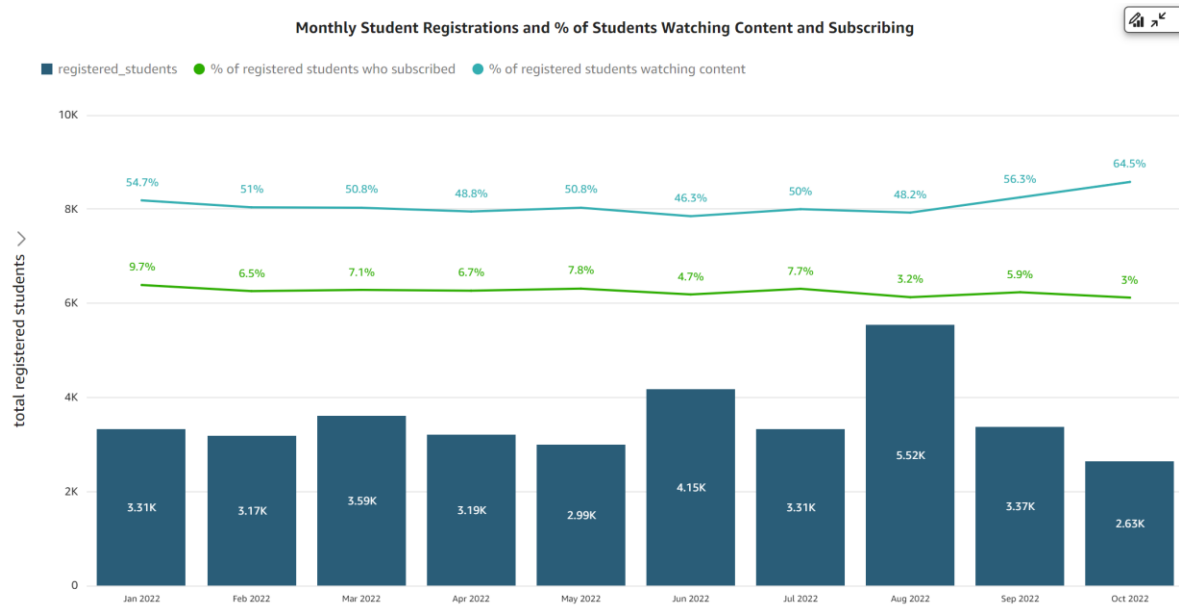
Part b. “Introduction to Data and Data Science” also has the highest student enrollment with 8.5K students, followed by “Introduction to Excel” with 2.4K students at an average rating of 4.78, and “Statistics” retaining third with 2K students and a rating of 4.81. Despite having a higher rating of 4.82, “Advanced SQL” ranks fourth in enrollment, with “Introduction to Excel” attracting approximately 1.5x more students, indicating a greater demand or preference for the latter course.



## Q2: How many students register monthly, and what fraction are also onboarded?

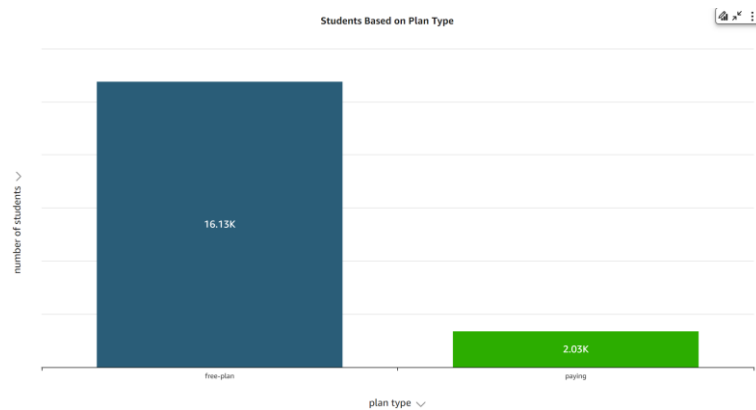
In January 2022, 3,312 students registered of which 321 students made a subscription purchase and were considered onboarded, approximating to 9.7% of the total registrations for that month. This is based on the definition of onboard students as those who have registered and made a purchase. Alternatively, using content watching as a measure of onboarding, 1,813 registered students were watching content and learning on the website and could also be considered as onboarded students, approximating 54.74% of total registrations for that month. The visual below summarizes monthly student registration, % of registered students purchasing subscriptions as well as % of registered students watching content and learning on the platform regardless of plan type (free-plan or paid subscription) (*Refer Query – A.2.1*)

Overall, out of 35,240 total registered students, ~52% (18,156) engage with content on the platform, and ~6% (2,135) have purchased a membership. (*Refer Query – A.2.2*)

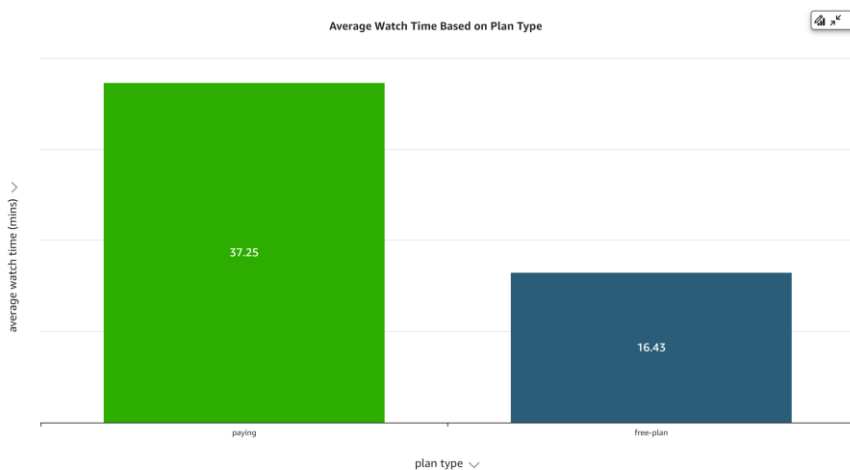


**Q3: How do students engage with the online platform (minutes and average minutes watched) based on type (free-plan or paying)?**

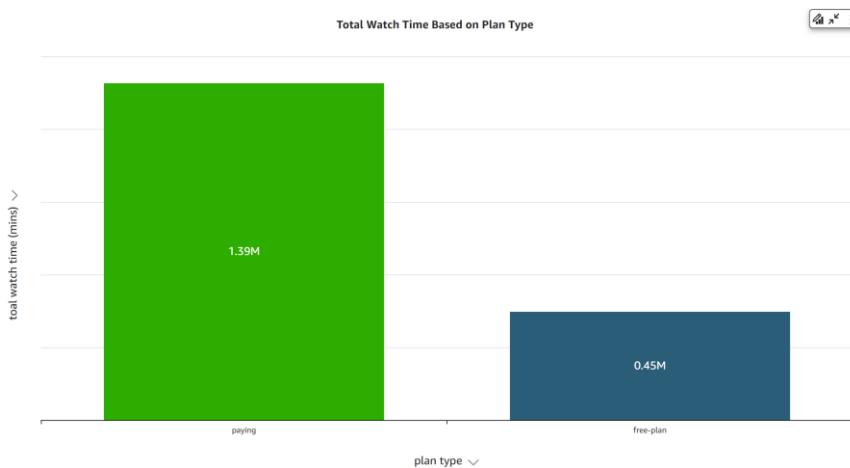
Of the 18,156 students engaged in learning, about 89% (16.13K) are on the free plan, while the remaining 11% (2.03K) are paying subscribers. (*Refer Query – A.3*)



Paying subscribers spend an average of 37.25 minutes per session, which is 56% higher than the 16.43 minutes watched on average by free-plan students. (*Refer Query – A.3*)

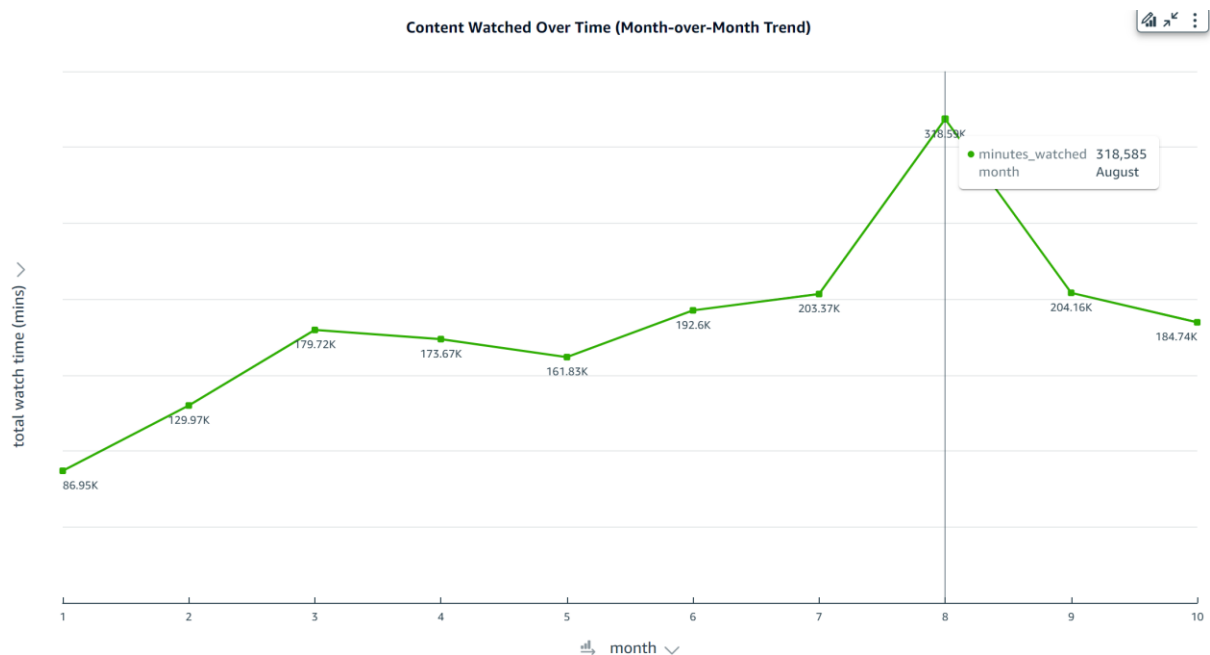


Overall, the paying students have watched a total of 1.39 million minutes, almost 3x more when compared to 0.45 million minutes by free plan students. (*Refer Query – A.3*)



#### Q4: Do students watch more content with time, and does it vary seasonally?

The trend between January and July shows a slight increase, with a notable peak in August at 0.32 million minutes watched, suggesting a seasonal peak. This trend levels off in the subsequent months. The visual representation of this trend can be seen below; however, the data from January to October 2022 is insufficient to confirm seasonality. While registrations also peaked in August (as seen in the chart from Q2 above), the available data does not establish a definitive link between the number of registrations and watch time (*Refer Query – A.4*)



**Q5: Which countries have the most students registered, and does this number scale proportionally with the number of minutes watched per country?**

(Refer Query – A.5)

With 6,933 students, India has the highest number of students registered and ~45% more registered users when compared to the United States with 4,768 students, but the United States leads with ~0.45 million minutes watched, ~150% higher than ~0.18 million minutes watched in India. The registration numbers do not scale proportionally with minutes watched when comparing countries, however, there is a strong positive correlation between the number of students registered and total minutes watched.

