

# Databases and Data Dashboards – A comparison of a locally hosted, free database and dashboard (MySQL/Excel) vs Amazon RDS/Amazon QuickSight dashboard (November 2017)

Jack K. Rasmus-Vorrath, *SMU MS Data Science student*, Nicole Shaver, *SMU MS Data Science student*

**Abstract**—Dashboards have become a popular way to display key success indicators at a glance to multiple users across an organization. This paper presents two possible methods for storing, retrieving, displaying and sharing data in a dashboard: 1) a MySQL database hosted on Amazon Web Services (AWS) for storing the data and Amazon QuickSight for creating and sharing the dashboard, and 2) local MySQL for storing the data, Microsoft Excel for creating the dashboard, and a Dropbox for sharing it. The two approaches are compared for cost, ease of creation and use, update speed, and shareability

**Index Terms**—data, dashboard, MySQL, Microsoft Excel, Amazon RDS, Amazon QuickSight.

## I. INTRODUCTION

DATA stored in a database is just data; sitting in a database, that data is of no value to an organization. The value of data lies in extracting it from the database and turning it into information that can be used to make decisions. Data dashboards have become a popular way to display key success indicators and metrics at a glance to multiple users across an organization, but there are many methods and options for storing, retrieving, displaying and sharing data in a dashboard.

In this research, we created a relational database of data in a stand-alone MySQL environment, and one hosted by Amazon RDS. We then created a dashboard in Amazon QuickSight and attempted to mimic that dashboard in Microsoft Excel. We then compared the cost, performance (i.e., speed to pull data and update the dashboard), features, ease of database and dashboard creation, and ability to share the dashboard across the members of our research group.

## II. PREVIOUS WORK

While there are numerous resources (papers, websites,

tutorials) on the advantages/disadvantages of freely creating and sharing dashboards with Excel and MySQL [1], as well as resources on how AWS QuickSight can be used to accomplish the same goals [2], there is relatively little discussion on how the two approaches compare. Most existing work evaluates pairing AWS with other third-party software tools like Tableau [3]. Issues of cost effectiveness are also of concern, particularly for small businesses and start-ups looking to cheaply build their own pipelines for streamlined in-house storage, retrieval, analysis, and presentation.

## III. RESEARCH METHODS

### A. Relational Database Creation and Normalization

The first step in the process was to create a relational database in our local copy of MySQL. The database was created by starting with a Kaggle dataset in CSV format on employee attrition [4]. From this flat file of data, the normalized database was created in a local copy of MySQL.

The import process involved resolving several type conversion and formatting issues, particularly with respect to datetime object types. A raw version of the database was first imported from the CSV file made available by Kaggle. After resolving formatting issues, this version was reverse engineered

This paper was submitted for MSDS7330 Term Project on November 20, 2017. We received no financial support and no sponsorship in completing this work.

Jack Rasmus-Vorrath is a student in the MS Data Science program at Southern Methodist University, Dallas, TX 75205 (email: [jrasmusvorrath@mail.smu.edu](mailto:jrasmusvorrath@mail.smu.edu)).

Nicole Shaver is a student in the MS Data Science program at Southern Methodist University, Dallas, TX 75205 (email: [nshaver@mail.smu.edu](mailto:nshaver@mail.smu.edu)).

to create an EER diagram, whose primary and foreign key values were accordingly modified and then synchronized with database. Synchronization autogenerated the code for proper normalized table creation, and this autogenerated code was then included in the final script.

The final script reproduces the whole importing, formatting, and normalizing process to facilitate reproducibility in the AWS/QuickSight setting. A raw flat table is created to prepare for import, making the necessary conversion and formatting adjustments. Tables are then created according to the normalized scheme using the EER autogenerated code. To avoid producing duplicate values, certain fields associable with individual employees had to be moved to the ‘status’ table, as there were several fields in the flat file which are updated for each employee on an annual basis (e.g., ‘Age’, ‘length\_of\_service’, ‘recorddate\_key’, and ‘STATUS\_YEAR’, all of which intuitively belong to the changing ‘status’ of individual employees). Duplicate values and processing errors also motivated the decision to link the ‘employee’ and ‘location’ tables by way of a concluding ‘UPDATE’ statement, which sets the autoincremented ‘location\_id’ primary key values from the ‘location’ table into the ‘employee’ table. Explicitly linking these two tables (via a ‘location\_id’ foreign key in the ‘employee’ table) to facilitate specific queries should be done at the user’s discretion, after all associated data have been unproblematically imported.

To replicate the local MySQL database for use with Amazon Web Services (AWS), we first created a database instance on AWS, and then set up the local MySQL database to connect to the AWS instance [6]. Once MySQL was connected to the AWS database, we ran the same code used to create the local version of the database, and the database was created in AWS.

### B. Amazon QuickSight Dashboard Creation

Several steps were required to create an Amazon QuickSight dashboard using data in an AWS instance of the database. We created a QuickSight account [7], adjusted the security settings of the AWS database instance to allow QuickSight to connect to our AWS database [8], imported the data from the AWS database (as separate tables) into QuickSight, and then used the QuickSight interface to join the tables together [9]. To create the dashboard in QuickSight, we dragged and dropped columns from the database into the dashboard interface, choosing the type of graph (i.e., line graph, scatter plot, pie chart) that was needed to best represent the data.

An overview of the dashboard creation process, from data source import to dashboard publication, is shown in Figure 1.

### C. Excel Data Dashboard Creation

To develop the Excel dashboard, we connected Excel to the local MySQL database, queried the local database to pull in the required data as a Table object, and then created a dashboard made up of pivot tables and charts to summarize the data and dynamically update upon new query results.

### D. Dashboard Shareability and Cost

To share the QuickSight dashboard among users, we simply clicked on the Share icon in QuickSight, and entered the email

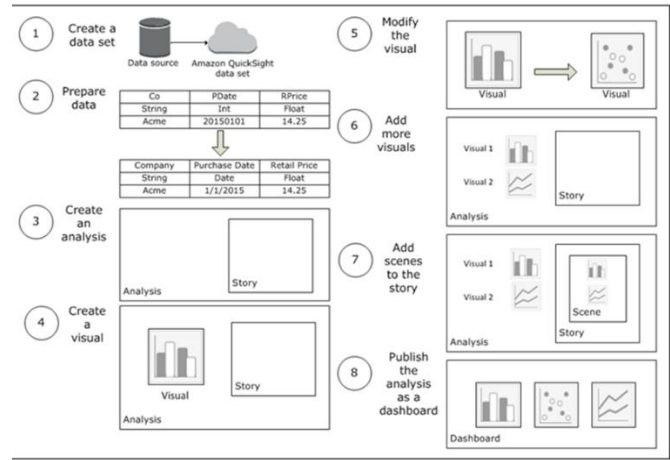


Fig. 1. Overview of data dashboard creation process [5]

address of the person we wanted to share the dashboard with. To share the Excel version of the dashboard, we created a shared folder in Dropbox, added the Excel Dashboard to that shared folder, and invited users to the Dropbox folder via email.

Finally, the cost of the Amazon database/QuickSight dashboard approach was compared with that of using local MySQL and Excel. We compared only the cost of dashboard creation and sharing, as comparing the cost of hosting and maintaining a database locally versus on a cloud involves multiple factors not explored in this study.

## IV. RESULTS

### A. Database Creation Comparison

The ease of database creation did not vary significantly between the local MySQL database and the AWS MySQL databases. In both cases, specialized knowledge of how to import the data records, assign the columns with proper formats, normalize the tables and create parent and foreign keys to link the tables was required. The script used to import the data is shown in the Index section below, and the EER diagram of the finalized database is shown in Figure 2. The import of the database from local MySQL to the AWS-hosted version was slightly more challenging for a novice AWS user, as one first had to set up an AWS database instance and set up the local MySQL workbench to connect to that database. However, once the connection was set up, the MySQL workbench worked seamlessly with AWS. At this phase of the dashboard creation, the two approaches did not differ significantly in difficulty.

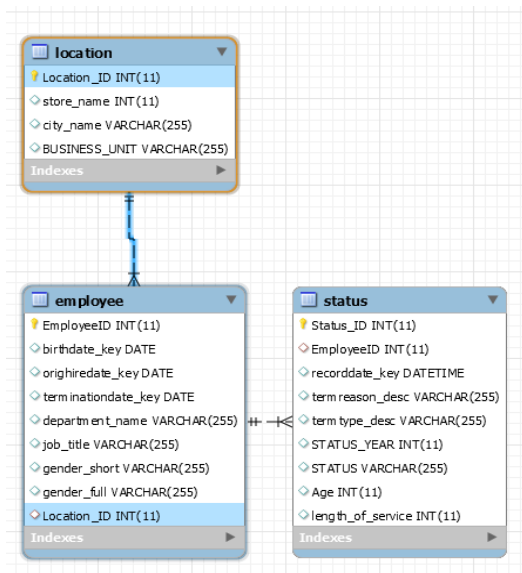


Fig. 2. EER diagram of database created from flat file of initial dataset.

### B. Dashboard Development Comparison

The creation of the dashboard in Amazon QuickSight took approximately 1 hour [10], following 2 hours of training and review of the documentation provided by Amazon [11]. The QuickSight platform offered multiple graphing options (pie charts, bar charts, stacked bar charts, scatter plots, line graphs, heat maps, pivot tables, and combination bar/line graphs), but those options are limited compared to the options offered in Excel, and the interface for creating the charts was not familiar. The customizability of the graphs was somewhat limited in QuickSight. For example, while one can change the range of an axis (e.g., from 0 to 100 to 0 to 30), one cannot specify tick mark size, the location of the tick marks, or whether one has minor and major tick marks or just major ones. Additionally, because the axes titles of the graphs are auto-generated, one cannot change the title without changing the name of the column in the data table in the data preview screen; one cannot change it directly in the interface that one uses to create the dashboard itself. Finally, when making line graphs, it was not possible to add points to the line, and the line of the graph was connected even when the data was not continuous. The dashboard created using Amazon QuickSight is shown in Figure 3.

Creation of an Excel dashboard requires varying amounts of training, depending on the user's familiarity with Excel. For a user with basic knowledge of how to manipulate charts, tables and graphs in Excel, translating that knowledge into creating automatically updated pivot charts and graphs from tables produced by database queries required approximately 1 hour of training. The basic idea of creating a dashboard from a flat file of fixed size was relatively simple, and was accomplished in approximately 1 hour after watching a simple 20-minute training video. Given a working knowledge of how to create and customize charts, it was easy to make changes such as



Fig. 3. Amazon QuickSight dashboard created from Amazon RDS version of database.

adjusting axis types and titles, changing axis ranges and tick mark settings, and changing data displays (to add or remove points from a line graph, for example). To create a dashboard that would query a database and automatically update, we used the database connection features available within Excel (Figure 4), importing the data from the MySQL database as a “Table” datatype, and creating the PivotTables and PivotCharts from that data Table [12].

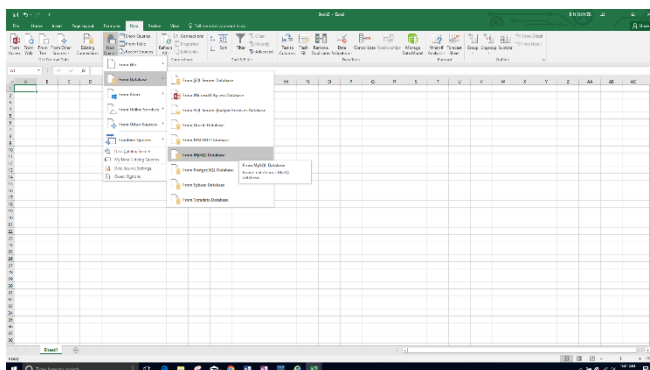


Fig. 4. Use of MS Excel database connection feature to pull data from local version of MySQL database into Excel as a Table Object.

Use of a Table Object in Excel means that the data will refresh when the query is rerun, and the PivotTable tables/charts created from that data table will automatically update to reflect the queried data. The dashboard created using this method is shown in Figure 5, and contains basic features of three typical dashboard types of charts (line, pie, bar) as well as a table.

The data dashboard created using MS Excel was very small when didn't have the data in it, and thus could be shared for free in Dropbox. In contrast, sharing dashboards in Amazon QuickSight costs \$9 per user per month.

While shareability is supposed to be one of the advantages of using Amazon QuickSight, we encountered a number of issues with sharing with one of the members of our group, as there

appeared to be some sort of conflict between their QuickSight account and Amazon RDS account that resulted in an unresolvable access error. Sharing via Dropbox was simple and we encountered no problems with this approach.

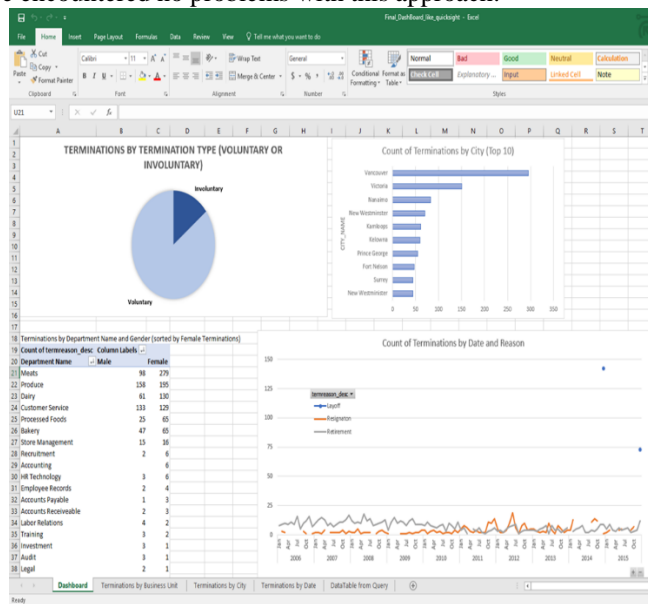


Fig. 5. Data dashboard created using MS Excel.

A table summarizing the cost, and the ease of database setup, dashboard creation and customizability, speed, and shareability is shown below.

	Amazon RDS MySQL Amazon QuickSight	Local MySQL/Excel
Dashboard Cost	QuickSight (includes creation and sharing): \$9/user/month (which includes 10GB of capacity/user)	Dashboard creation-Free (assuming MS Excel installed on computer)  Dashboard sharing-Free on <a href="#">Dropbox</a> (for up to 2GB; \$10/user/month for 100GB). Dashboard file (not including dataset) is 0.001 GB
Ease of database setup	Medium-required set up of AWS database instance and connection of MySQL to that instance	Easy-no special set up required on local MySQL
Ease of Database/table creation	No difference between local and AWS once connection is established	No difference between local and AWS once connection is established
Ease of Dashboard creation	2 hours of training/documentation review  0.5 hours to set up Amazon RDS to connect to QuickSight (required adjustment of security settings)  2 hours to create QuickSight Dashboard	1 hour of training/documentation review  0.5 hours to import data into Excel  1 hour to create Excel dashboard
Dashboard customizability	Medium-cannot edit titles without editing column names, limited axis/scale customization, limited chart types	High-all of the options available in Excel
Speed to update/Refresh	Instantaneous	30 seconds
Shareability	Available on any computer or mobile device, anytime  Theoretically easy but issues with sharing with some users—account would not allow access	Shared by using Dropbox to secure folder. Only viewable on computer (not mobile device) with Excel installed  Different versions of Excel had varying access to updating the pivot tables/charts, but all versions could view the tables/charts

## V. CONCLUSION

In this paper we compared Microsoft Excel and Amazon QuickSight as two options for sharing and displaying data in a data dashboard. While Amazon QuickSight offered better shareability across users and platforms, and was not hampered by such issues as differing versions of Excel, the sharing was not seamless (we were not able to share the dashboard with all team members due to AWS account conflict issues) and comes at a significant cost (\$9/month/user). Additionally, while Amazon QuickSight is advertised as easy to set up and use, pulling the data from an AWS data source proved more complicated than expected, and the interface for creating graphics was foreign, and not as customizable as expected. In contrast, after setting up the system for pulling data from a database into Excel, the creation of dynamically updating charts and graphs was trivially easy for someone with a basic working knowledge of Excel charts.

For companies without major budget constraints, already using AWS products, who require dashboards to be shared across multiple platforms (Mac and Windows, mobile and desktop/laptop), and who do not require highly customized charts and graphs, Amazon QuickSight is an excellent solution. However, for companies requiring more customized charts, who do not need mobile/cross-platform shareability of their dashboard, and who are looking for a less expensive solution, Excel is a powerful data manipulation/chart creation tool that can be used to make a data dashboard with minimal additional training.

## APPENDIX

Code for importing csv datafile and creating tables in database can be found at the following GitHub link:

[https://github.com/JackKRasmus-Vorrath/Databases\\_and\\_Dashboards\\_-\\_MySQL-Excel\\_vs.\\_Amazon\\_RDS-QuickSight/blob/master/Databases\\_mydb\\_import-and-process\\_final.sql](https://github.com/JackKRasmus-Vorrath/Databases_and_Dashboards_-_MySQL-Excel_vs._Amazon_RDS-QuickSight/blob/master/Databases_mydb_import-and-process_final.sql)

Project repository:

[https://github.com/JackKRasmus-Vorrath/Databases\\_and\\_Dashboards\\_-\\_MySQL-Excel\\_vs.\\_Amazon\\_RDS-QuickSight](https://github.com/JackKRasmus-Vorrath/Databases_and_Dashboards_-_MySQL-Excel_vs._Amazon_RDS-QuickSight)

## ACKNOWLEDGMENT

J. K. Rasmus-Vorrath thanks Ian Kinskey for comments and suggestions on developing the MySQL database. N. Shaver thanks her family for their support while she completed this project.

## REFERENCES

- [1] "Interactive dashboard from MS SQL Server, MySQL, Excel", YouTube, 2017. [Online]. Available: <https://www.youtube.com/watch?v=Mm-iFlzT2-o>. [Accessed: 17- Nov- 2017].
- [2] R. Natarajan, "Visualize MySQL data with AWS QuickSight", Linked-In, 2017. [Online]. Available:

<https://www.linkedin.com/pulse/visualize-mysql-data-aws-quicksight-rahul-natarajan/>. [Accessed: 17- Nov- 2017].

[3] "Tableau Server on the AWS Cloud", *S3.amazonaws.com*, 2017. [Online]. Available:

<https://s3.amazonaws.com/quickstart-reference/tableau/server/latest/doc/tableau-server-on-the-aws-cloud.pdf>. [Accessed: 17- Nov- 2017].

[4] "Employee Attrition | Kaggle", *Kaggle.com*, 2017. [Online]. Available:

<https://www.kaggle.com/HRAnalyticRepository/employee-attrition-data>. [Accessed: 17- Nov- 2017].

[5] "Welcome to Amazon QuickSight! - Amazon QuickSight", *Docs.aws.amazon.com*, 2017. [Online].

Available:

<http://docs.aws.amazon.com/quicksight/latest/user/welcome.html>. [Accessed: 20- Nov- 2017].

[6] "Connect to RDS with MySQL Workbench", *Amazon Web Services, Inc.*, 2017. [Online]. Available:

<https://aws.amazon.com/premiumsupport/knowledge-center/connect-rds-mysql-workbench/>. [Accessed: 17- Nov- 2017].

[7] "Sign Up for Amazon QuickSight with an Existing AWS Account - Amazon QuickSight", *Docs.aws.amazon.com*, 2017. [Online]. Available:

<http://docs.aws.amazon.com/quicksight/latest/user/sign-up-existing.html>. [Accessed: 17- Nov- 2017].

[8] "Authorizing Connections from Amazon QuickSight to Amazon RDS DB Instances - Amazon

QuickSight", *Docs.aws.amazon.com*, 2017. [Online].

Available:

<http://docs.aws.amazon.com/quicksight/latest/user/enabling-access-rds.html>. [Accessed: 17- Nov- 2017].

[9] "Joining Tables - Amazon

QuickSight", *Docs.aws.amazon.com*, 2017. [Online].

Available:

<http://docs.aws.amazon.com/quicksight/latest/user/joining-tables.html>. [Accessed: 17- Nov- 2017].

[10] "Working with Amazon QuickSight Visuals - Amazon QuickSight", *Docs.aws.amazon.com*, 2017. [Online].

Available:

<http://docs.aws.amazon.com/quicksight/latest/user/working-with-visuals.html>. [Accessed: 17- Nov- 2017].

[11] "Tutorial: Modify Visuals - Amazon

QuickSight", *Docs.aws.amazon.com*, 2017. [Online].

Available:

<http://docs.aws.amazon.com/quicksight/latest/user/example-modify-visuals.html>. [Accessed: 17- Nov- 2017].

[12] "MySQL :: MySQL for Excel Guide :: 5.6 Creating PivotTables", *Dev.mysql.com*, 2017. [Online]. Available:

<https://dev.mysql.com/doc/mysql-for-excel/en/mysql-for-excel-import-pivottables.html>. [Accessed: 17- Nov- 2017].