

## **DATA WAREHOUSING PROJECT: Lyft**



### **SUBMITTED BY:**

- **AYUSHI SHRIVASTAVA – tf2272**
- **SHUBHANGI JAIN – qt9379**

1. The objective of this group project is to develop Data Lake architecture for a Mid-Size business in an industry of your choosing (Excluding retail and hotel).

Our objective is to develop a Data Lake Architecture for LYFT.

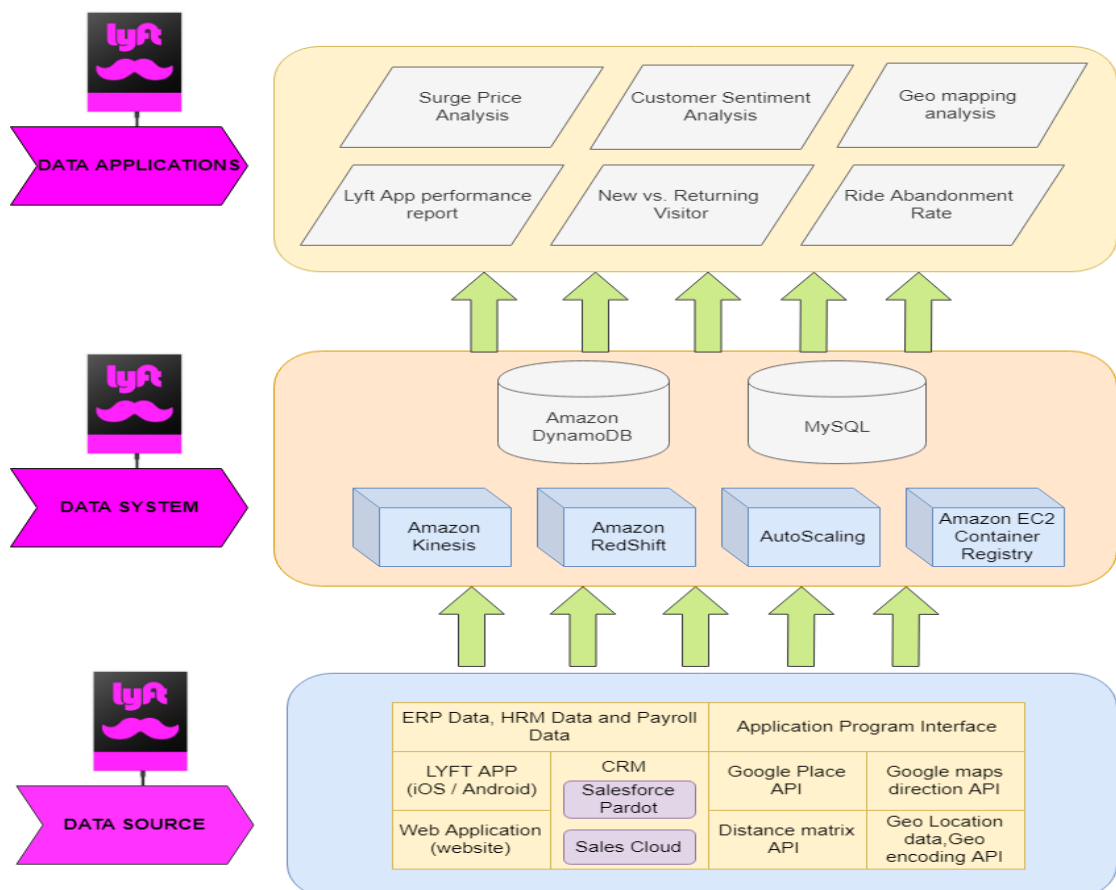
**Industry:** Transportation

**Mid-Size Business:** LYFT

2. Develop and submit a report describing the Data Lake Architecture for the Mid-Size company.

Lyft is using mainly two databases to store the data it retrieves from the app. The first one is Amazon DynamoDB, Amazon Kinesis, Amazon RedShift, Auto Scaling, Amazon EC2 Container Registry and the other one is MySQL.

We assume that the unstructured data will go in Amazon DynamoDB, Amazon Kinesis, Amazon RedShift, Auto Scaling, Amazon EC2 Container Registry and the structured data will be stored in MySQL.



## Data Source

### Structured Data Types : ( MySQL Database)

- ✚ **Transactional Data:** This data pertains to the transactional data and is the structured component of Lyft's database. The app also records the driver details, rider details, driver car details, location of rider details, trip duration details, trip cost details, rides payment data and enables secure transactions through an encryption.
- ✚ **User (Rider/Driver) Data:** Personal user information retrieved by user's Lyft account and information about Cars, Vehicle IDs, and the type of vehicles is stored in this type of data. Riders are the users who want to use Lyft's services to get a ride from their source to the destination. Lyft asks for specific details pertaining to **User** while they sign-up like First Name, Last Name, Address, Contact Details, Payment Method etc. It also maintains details of **Driver** fetched through Lyft Driver App wherein it has information pertaining to Driver's vehicle (Year, Make, Model, Color, Number of Doors, and Number of Seat Belts) and Driver's personal details like SSN, Name, Address, Age etc. SSN is the important requirement to be able to register as a Driver in Lyft.
- ✚ **CRM Systems Data:** Lyft is using **Salesforce Pardot and Sales Cloud** to identify the best prospects, convert more leads and drive customer loyalty. Lyft mainly uses Salesforce CRM software for their business model. Lyft Business will be able to drive customer engagement and loyalty along with other tools such as Sales Cloud. They know how important it is to serve their customers, keep track of every transaction and interaction, and use every piece of information to create new possible business opportunities. This is one of the reasons Lyft was able to reach the milestone of 1 million riders within a year of its launch in San Francisco.
- ✚ **Human Resource Management Data:** The human resource activities involve managing employee relations, managing employee payroll, benefits and compensations, appraisals, recruiting the employees, dealing with performance issues and discrimination. Human Resource representatives at Lyft take the data from our recent diversity to analyse the growing need for more policies to support its employees. These new policies demonstrate our commitment to ensuring that the employees who work here are treated with the utmost respect by their five thousand colleagues.
- ✚ **Payroll Data:** Lyft using other internal data, performed recalculations using Buzz feed's expense formula and processed earnings numbers. But even after all that convoluted math. They pay the portion typically reserved for employers along with standard payroll taxes.

### Unstructured Data Examples: (AmazonDynamoDB Database):

- ✚ **Amazon DynamoDB:** It is used for multiple data stores, including a ride-tracking system that stores GPS coordinates for all rides. Lyft's click stream and web logs data is used for understanding customer preferences and improving their experience.

🚦 **Amazon RedShift**: It is used to gain customer insights that power the company's shared-ride product, Lyft Line.

🚦 **Auto Scaling**: To track the number of bookings being done in a particular area. Lyft generates a different data pattern that contains frequency of bookings at certain time of the day. These patterns are used to predict possible surges in bookings in an area so that Lyft can allocate adequate amount of drivers. It is used to manage up to eight times more riders during peak hours.

🚦 **Amazon Kinesis**: It is used to funnel production events through the system, and leverages the scalability of Amazon DynamoDB for multiple data stores, including a ride-tracking system that stores GPS coordinates for all rides.

🚦 **Amazon EC2 Container Registry**: As part of their micro services infrastructure on AWS, Lyft relies on Amazon EC2 Container Registry (ECR) to durably store container images for their applications and reliably deliver these images to downstream test and deployment systems.

- **Lyft** also takes the help of APIs such as Google Maps Direction, Google Place, Distance Matrix and Google Map Geo. These API provide Lyft with following functionalities :-
- Finding the exact location of the passenger/rider.
- Selecting the route with minimum distance and traffic.
- Showing the estimated cost of the trip based on the distance.
- Showing the rider details sharing the ride in "Shared" category.
- Displaying the Estimated Time of reaching the destination considering the traffic.
- Displaying different choices of car to select for the ride from following categories: Economy, Extra Seats and Luxury.

**Location Coordinates and Booking Confirmation** – It encompasses or tracks Longitudinal and Latitudinal Co-ordinates of the Location from where users are booking Lyft. This information is used to set the route for the driver and it books a ride of a driver which is the nearest to the rider's location. As the passenger confirms a cab request, a notification is sent out to the closest drivers. After a ride is booked, rider is sent the trip details including current location of the driver, estimated time for the arrival of driver to the location, car details (Number Plate, Color, Make, and Driver's Name). To request a cab, the app utilizes Google Maps to allow users to even drop a pin on the map giving exact location if they couldn't find it in the search. The distance, ETA can be easily returned using Google Map APIs.

**Trip data**: Lyft records data points based on the time of the trip, and whether the passengers reached on time. It keeps track of the trip at very granular level. If rider takes time to board the ride and makes the driver to wait, the trip cost is revised at the end of the trip. Rider also has the option to cancel the ride if he/she finds that driver is going to take longer to arrive or the trip is costing too much or faces any arguments with the driver etc. Rider can also ask for the refund of the cancellation charge if he/she finds it to be charged unfairly by stating the reason. All this information regarding the trip is encapsulated in the trip data.

### **Data System (MySQL, Amazon DynamoDB)**

Lyft's data system consists of MySQL and Amazon DynamoDB, Amazon Kinesis, Amazon RedShift, Auto Scaling, Amazon EC2 Container Registry for organizing the data. Both of these databases are used for separate scenarios.

MySQL is used to organize and store permanent data like account information of both passengers and drivers. Whereas the Amazon DynamoDB is used for real-time applications like storing, queuing and processing real-time bookings of thousands of users in a minute.

The summary of every ride is stored in the data system which is essential for Lyft to find booking behaviour and hot spots in any area. By analysing the data, Lyft can calculate the demands of rides at any particular hour at any location. This way Lyft is able to allocate the required number of drivers as well as impose a surge price which helps them to get maximum profit.

Lyft initiated testing the app with a small group of users, and the combination of qualitative and quantitative analysis helped them refine the experience before they launched more widely. Quantitative Research helped them find areas of opportunity: for example, they saw that users had difficulty finding other ride types in an earlier version of the app. Qualitative Research helped them more deeply understand the issue, and then validate solutions that they should build and experiment with. Lyft product team talked to users in-person to get feedback on why the previous design made it challenging to discover the full list of ride types and to compare that to new designs that we were proposing. Users' reactions to the new designs narrowed our list of potential solutions before they built one and again returned to the quantitative side by running a large-scale experiment. The results from that experiment gave them the final answer on whether our solution effectively solved the initial problem. By working closely with each of their customers, Lyft Business has discovered how complex the customer lifecycle is and the varying challenges that each industry has when it comes to transportation.

When there is a sudden rise in the frequency of Lyft booking in a particular area at a particular time, the Surge price analysis creates a pattern of this data and predicts the next possible occasion for sudden increase in the frequency of Lyft bookings. It uses products such as Auto Scaling to manage up to eight times more riders during peak.

On the other hand the business intelligence analysis is a tool that is extensively used to support decision-making. Customer Sentiments analysis is done based on the customer hits to the app, returning customers, dissatisfied customers, potential customers etc. Every customer's profile who has signed up for Lyft is analysed based on many metrics like Age, Frequency of rides, gender, monthly expenditure, whether they have enrolled for Lyft's rewards or monthly plans etc. Based on the analysis, it creates customised customer loyalty programs, promotions and rewards. The Loyalty program is an essential tool to improve customer retention. Therefore every user gets the most appropriate discount for themselves. This way Lyft is able to keep their customers happy as well as manage to have a Competitive Edge over its competitors.

### **3. a) Short Overview of Business:**

Lyft is the fastest growing rideshare company in the United States that enables users to book rides based on their location. Lyft has an interactive application where in a user can

sign up to be a driver or just a rider. It is available in more than 200 cities in the United States and 9 cities in Canada, facilitating 14 million rides per month. It was launched in May 2012 and has achieved their first milestone of 1 million rides in a year. Lyft is the fastest growing company of its kind in the United States, reporting 200 percent year-over-year growth in 2015 for both rides and revenue. Lyft, with a mission of reconnecting people and communities through better transportation, is working with General Motors to create a network of on-demand autonomous vehicles and has partnered with local market leaders Didi Kuaidi, Ola, and GrabTaxi to launch its international expansion in Asia.



### **3.b) Transactional data in relational databases (Data sources and systems corresponding to each value chain activity, schema not required)**

- ✚ **Inbound Logistics:** Lyft, the fastest growing company in the United States, does not own the vehicles it uses to serve customers. The vehicles are owned by Lyft drivers or Lyft offers a vehicle rental program for drivers called Express Drive and the cost of the rentals is around \$180 – \$240/week. Lyft drivers and Lyft users need to possess smartphones to use the Lyft app or mobile website to be able to use the services.



- ✚ **Operations:** Lyft operates in more than 200 cities in the United States and 9 cities in Canada. Lyft services involve the following stages:
  - **Requesting the ride:** Customers can use the Lyft app to tap each ride option to see wait time, size, and price. Customers can enter their pickup location and tap-request for their driver to arrive in minutes.
  - **Ride:** Customers are able to see the contact information of their driver, as well as, vehicle details on the app. This ensures that customers get into the right car.
  - **Payment:** When they reach their destination, customers simply hop out of the car and give the rating to their driver. Customers don't need to have cash or credit cards with them, it automatically charges from the credit card.
- ✚ **Outbound Logistics:** Outbound logistics in value chain analysis is associated with warehousing and distribution of products. Outbound logistics does not fully apply to this service sector such as ride-hailing due to inseparability of delivery and consumption of services. As discussed above, Lyft operates in more than 200 cities in the United States and 9 cities in Canada and internet-based nature of its business model is a major source of value for the company.
- ✚ **Services:** Customer Satisfaction is number one priority for Lyft. Good customer service is effectively integrated into Lyft business model. Riders rate their drivers on the scale of 1 to 5 after each ride and drivers with the average low riding scores are promptly deactivated. Drivers therefore have an incentive to maintain a clean car and provide good service. Customers can complain if their driver took a longer route to reach the destination or driver didn't reach on time for the pickup, etc. Lyft is assessed to be prompt in issuing refunds in such situations.

Lyft offers five types of rides within the app:

- **Shared Ride-** which is not available in all cities, is the cheapest option and will match passengers with other riders if they are going in the same direction
  - **Lyft-** it is the basic and most popular offering that matches passengers with nearby drivers
  - **Lyft XL-** it matches passengers with a vehicle that can seat at least 6 passengers
  - **Lux-** it matches passengers with a luxury vehicle that seats at least 4 passengers
  - **Lux Black-** it matches passengers with a luxury black exterior vehicle ride that seats at least four passengers
  - **Lux Black XL-** it matches passengers with a black exterior luxury car that seats up to 6
- ✚ **Marketing and Sales:** It mainly relies on social media marketing and word-of-mouth marketing to communicate its marketing message to the target customers. Moreover, Lyft uses media advertising, sales promotions and public relations as part of its marketing strategy. Convenient payment method is a major source of value addition for the company. Lyft automatically charges the credit card registered with the app. Therefore, Customers don't have to have cash or credit card when they reach destinations.

### **3.b.ii) Big data from Social and/or Sensor or other sources.**

The Lyft sensor will give you time and price estimates for all available Lyft products at the given start latitude and start longitude. The ATTRIBUTES are used to provide extra information about products, such as vehicle capacity and fare rates. If end latitude and end longitudes are specified, a price estimate will also be provided. One sensor will be

created for each product at the given start location, for pickup time. A second sensor for each product, for estimated price, will be created if a destination is specified. The sensor is powered by the official Lyft API. The various parameters captured by sensors are used for analysis such as the number of customers, drivers with vehicles at a particular location identified by Latitude and Longitudinal attributes. The impact of weather attributes and Traffic conditions on Lyft Sales.

ETL is a process to extract data from various raw events, transform them for analysis and load the derived data into a query able data store. Data engineers and scientists at Lyft build various ETL pipelines which run at a different set schedule to gain insight on topics ranging from the current ridesharing market to the experiences for driver/passenger, etc. A reliable, efficient, and trustworthy workflow management system is crucial to make sure these pipelines run successfully and deliver the data on its set schedule. Lyft is the very first Airflow adopter in production since the project was open sourced around three years ago. Today, Airflow has become one of the most important pieces of infrastructure at Lyft which serves various use cases: from powering executive dashboards to metrics aggregation, to derived data generation, to machine learning feature computation, etc.

Similarly for traffic, based on the data mining process, we can transform traffic scores to different types and ranges of traffic to incorporate this into a data warehouse. The data on weather can be transformed during the ETL process to incorporate weather conditions in a certain range ranging from extreme cold, precipitation - snow, precipitation - rain, cool, warm to extreme hot. Data from various social media platforms is collected as raw data and sentiments can be extracted which gives an overall perspective how the customer feel about ride sharing services.

Market research data is an important factor which helps to analyze the competitor and generate strategies.

### **3.c) Enterprise Data Warehouse Bus Architecture for the Mid-Size Business**

	Driver	Customer	Vehicle	Date	Time	Location/ Tracking	Employee	Refund	Payment	Customer Feedback	Promotions
Transaction	X	X	X	X	X	X	X	X	X	X	X
Cancellations	X	X	X	X	X	X	X	X	X	X	
Legal	X	X	X	X	X	X	X			X	
HR Recruitment	X	X		X	X	X	X			X	
Marketing/ Loyalty programs		X		X	X	X	X	X		X	X
Surge Pricing	X	X		X	X	X	X	X	X	X	X
Service	X	X	X	X	X	X	X	X	X	X	X

The enterprise data warehouse bus matrix is the essential tool for designing and communicating the enterprise data warehouse bus architecture. The rows of the matrix are business processes and the columns are conformed dimensions. The shaded cells of the matrix indicate whether a dimension is associated with a given business process. The design team scans each row to test whether a candidate dimension is well-defined for the business process and also scans each column to see where a dimension should be



conformed across multiple business processes. Besides the technical design considerations, the bus matrix is used as input to prioritize DW/BI projects with business management as teams should implement one row of the matrix at a time.

### **3.d) Calculations showing the size of Data Warehouse**

#### **i) Make Assumptions on Number of Records for dimension tables and Number of Records per quarter on fact tables**

Our rough estimates for the rows generated per quarter for dimensions are as follows:

Dimension Table	Number of Records
Vehicles	1400000
Customer Demographics	4
Driver Demographics	4
Customer	18600000
Traffic	10
Time	80000
Driver	1400000
Tracking	300000
Weather	10
UTC Time	80000
Date	90

- Lyft has 18.6 million customers, 1.6 million drivers and does an average of **1 million rides per day**
- Age and Income level have 4 categories, Weather has 10 categories
- **Age** – 18-25, 26-35, 36-45, 46-55
- **Income** - <\$15000, \$15001-\$30000, \$30001-\$45000, >\$45000

- Number of rows in dimension tables = 50 million
- Number of records per quarter in fact table 1 million \* 90 = 90 million / quarter
- Total number of records / quarters = **140 million.**

### 3.e) . Analysis of relational vs. non-relational storage options

#### i) For each data source identified in (b) discuss the best way to store the data (relational vs. non-relational) and the underlying reasoning.

As discussed above the data is structured, so storing it in a relational database is a good choice. By doing so it provides a layer of protection, preserves the integrity of the data and makes it easier to meet compliance. Relational databases support an important concept of dynamic views. In this database, a view is not a part of the physical schema, it is dynamic. Hence changing the data in a table alters the data depicted by the view. Views can subset data, join and simplify multiple relations, dynamically hide the complexity in the data, and reduce data storage requirements. Moreover, Relational databases provide excellent security. They support access permissions which allow the database administrator to implement need-based permissions to the access of data in database tables. Relational databases support the concept of users and user rights, thus meeting the security needs of databases. The other important advantages of relational databases include their performance, power, and support to new hardware technologies, as also flexibility and a capacity to meet all types of data needs. These databases are scalable and provide support for the implementation of distributed systems. The App records each transaction, and connects through a secure API to the customers payment vendor. Approximately 75% of the amount is transferred to the driver. All this data about customers, the payments made by customers and all the transactions among them is stored in a relational format. Thus, **Mobile Applications/ Web Application and Transactional data is stored in a relational database (MySQL).**

There is huge amount of real time data, which means unstructured data is being captured at every instance and to handle such huge amount of data non-relational database is considered as the best. Data is stored in Amazon DynamoDB, because it is good at indexing geolocation. There are lot of activities in a single transaction. For example, if a customer request for a ride then the request is transferred to the nearby driver. After the driver accepts the ride, the customer can see the ETA (Estimated time arrival) if it is less than 5 mins, customers wait for the ride. If the ETA is longer than 5 mins or even more then customers cancel the ride. Therefore, because so much of back-and-forth read-write activity is needed in a single transaction, a relational database could never keep up with the speed and scaling necessary. Additionally, non-relational databases trades consistency with scalability. Thus, **Big Data from Sensor, Machine generated and geolocation data is stored in non-relational database (Amazon DynamoDB)**

Assuming 2KB per row, size of data warehouse = 2KB \* 140 million = 280 GB / quarter = 1.1 TB / Year

- **Cost of storing 1.1 TB in a relational database:**

(<https://azure.microsoft.com/en-us/pricing/details/sql-database/managed/>)

\$0.0580 per month/ GB.  $0.0575 * 1100 * 12 = \$759$

Cost of SQL 2017 Enterprise 2-pack of Core Licenses = \$14,499

Total cost = \$15258

➤ **Cost of storing 1.1 TB in a non - relational database:**

(<https://aws.amazon.com/dynamodb/pricing/>)

\$0.25 per GB per month  $0.25 * 1100 * 12 = \$3,300$

**ii) Present 2 options for big data storage and cost estimates.**

**Big data storage options:**

➤ Amazon AWS([https://aws.amazon.com/pricing/?nc2=h\\_ql\\_pr&awsm=q1-3](https://aws.amazon.com/pricing/?nc2=h_ql_pr&awsm=q1-3))

Cost of storing 1.1 TB per year =  $0.0277 * 1100 * 12 = \$365$

➤ Microsoft AzureHD(<https://azure.microsoft.com/en-us/pricing/details/hdinsight/>)

$0.0184 * 1100 * 12 = \$243$

**3.f) Data Lake Architecture that describes how/where Data Warehouse and Big Data are stored and how they can be integrated for decision support. The architecture should include storage, processing and reporting layers for different types of data.**

A data lake is a storage repository that holds a vast amount of raw data in its native format, including structured, semi-structured, and unstructured data. The data structure and requirements are not defined until the data is needed.

The data lake architecture identifies the various layers for storage and reporting for different types of data. The data from sensors, Google maps for traffic, Google maps for locations and weather conditions is in the source data. The source data is stored in the Data Lake depicted as Amazon DynamoDB below. The transactional data, based on payments, involving two parties, drivers and customers is stored in the MySQL database; this contains all the details about transactions, and profit per ride. From Amazon DynamoDB the data is transformed using clustering, mining and simpler data can be transformed using ETL processes (example weather ranges, Surge Pricing when prices are surging, you'll see a multiplier to the standard rates on the map. For example, you might see surge at 1.2x or 2.2x. This is how much your base fare will be multiplied by, so a fare that is usually \$12 would be \$22 when it's at 1.8x Surge. Lyft's fee percentage does not change during surge pricing. Surge pricing on the map shows the price that will apply to riders using the Lyft app in that area. The rider's location determines the amount of surge pricing on a trip, not the driver's location.)

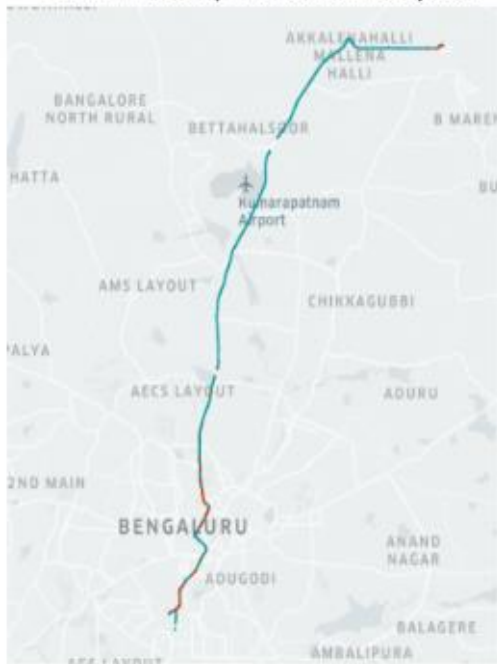
There are several important applications of transforming this data and using it for business decisions. We will be using the MySQL data for sales analysis, however integrating this with different types of sensor data such as locations, weather, and traffic scores gives us the advantage of predictive modelling as well. By using sensor data and clustering different traffic conditions, and locations, we can arrive at traffic ranges, and locations ie hotspots. The two aspects of our applications, the first one being BI and reporting, Sales and loyalty will be part of that data, as sales data depends on daily transactions and loyalty data depends on the amount spent by a customer and both these types of data are stored in the MySQL database in a relational format.

**3.g)**

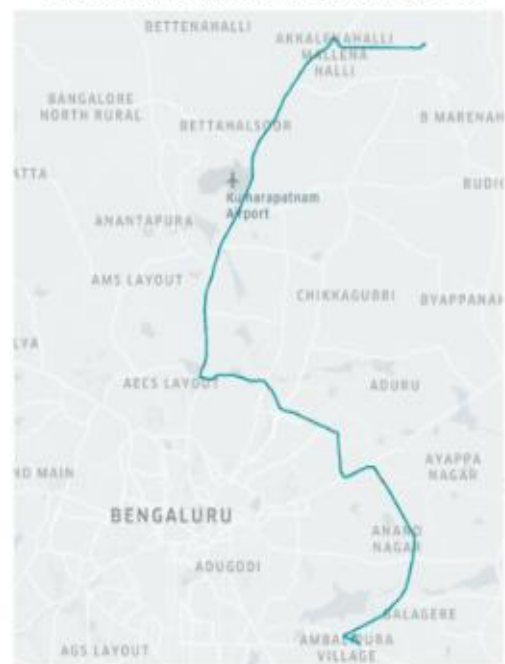
- i) Sample Reports or Schema that show how Big Data and Data Warehouse can be integrated to solve specific business decision problems relevant to the company/industry

Report 1	
Report Name	GPS Spoofing Detection
Report Description	<p>Fraudsters use GPS spoofing apps/techniques to create fake locations on a different phone in order to simulate a real trip in the hopes of getting paid through a driver account. The standard technique involves a fraudster creating a new rider account, adding a stolen credit card, and using that account to pay for a spoofed/fake trip from their driver account. The credit card is cashed out and payment goes to the fraudulent driver account.</p> <p>Sometimes fraudsters create multiple spoofed/fake trips to boost the total number of trips, so they can earn an incentive bonus from Lyft.</p> <p>Solution:</p> <p>By developing a speed profile for roads in an area, we can figure out which trips deviate substantially from their standard profile trends, likely indicating GPS spoofing.</p>
OLAP Report Column Labels	Date hierarchy collapsed with Time
OLAP Report Row Labels	Hotspots
Report Filters	Traffic Condition
OLAP Report Cell Values	Average Trip time

14.5% of the trip has abnormal speed



0.2% of the trip has abnormal speed



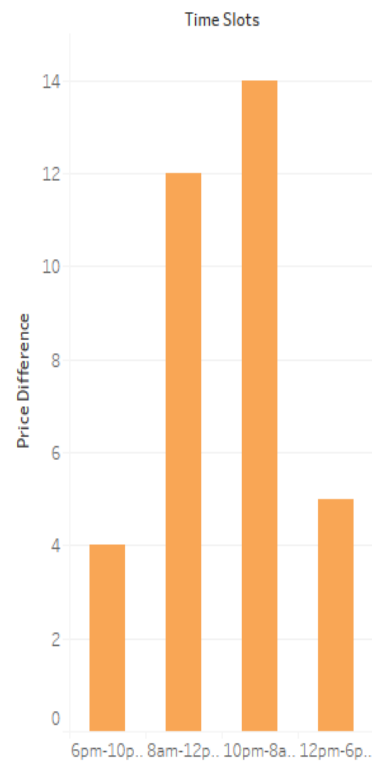
■ Abnormal speed  
■ Normal speed

Report 2	
Report Name	Difference between rides shared and normal Lyft during time of the day
Report Description	<p>This report will help us to obtain a relation on customer ride preference vs the time of the day</p> <p>For example during 10:00 pm the cost of shared Lyft is \$5 while normal Lyft is \$6.8 so which ride would the customers prefer?</p> <p>Benefits: The company would have the detailed customer churn model of customer preference vs \$price difference and set the price of the rides on the basis of business development goals.</p>
OLAP Report Column Labels	Date hierarchy collapsed with Time
OLAP Report Row Labels	Hotspots
Report Filters	Type of ride
OLAP Report Cell Values	Average Cost per trip

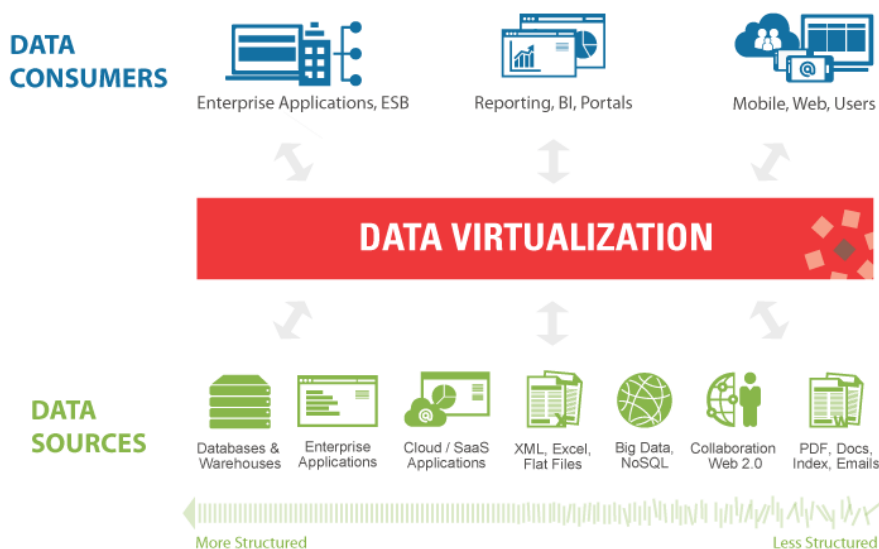
Time slots vs Trips



Time slot vs Price Difference



**3.h) Data Virtualization: What is data virtualization? How can the company can benefit from data virtualization? How does it fit into the organizations data infrastructure?**





## **What is data virtualization?**

Data Virtualization (DV) is unlike traditional Data Integration, wherein change is made on multiple layers Data Virtualization makes change easy for the business as new requirements and sources can be integrated and changed rapidly according to agile methodologies.

Data Virtualization enables distributed databases, as well as multiple heterogeneous data stores, to be accessed and viewed as a single database. Rather than physically performing ETL on data with transformation engines, Data Virtualization servers perform ETL processes such as data extract, transform and integrate virtually.

Data virtualization is a modernized approach to data management systems that allows application/consumers/end users to retrieve and manipulate data without requiring technical details about the data, such as how it is formatted at source, or where it is physically located, and can provide a single customer view of the overall data.

### **Advantages**

- Operation Costs
- Time to solution
- Security (abstraction)
- Data quality (Error rate)
- Compliance
- Zero Replication
- Agility

### **How does it fit into the organizations data infrastructure?**

- Lyft has a business model of payment via cards as soon as a ride is booked here data and credit card confidentiality is pivotal and hence virtualization comes into picture to provide security for a customer base of 23 million
- ETL is constant in Lyft apps ranging from credit card transactions to geo locations to customer support. Hence to support such kind of data flow we require fast process as well as reliability.
- Multiple transactions/rides are performed by a single driver/rider, there should be no redundancy.
- ETL cost is huge so is the customer base hence virtualizations is a more efficient solution than traditional databases.
- Lyft follows Agile data services provisioning - Data virtualization promotes the API economy. Any primary, derived, integrated or virtual data source can be made accessible in a different format or protocol than the original, with controlled access in a matter of minutes.
- Lastly fraudsters/hackers are constant trying to acquire customer credit card details hence the level of abstraction provided by virtualization is must.

### **References:**

1. <https://aws.amazon.com/solutions/case-studies/lyft/>
2. <https://aws.amazon.com/solutions/case-studies/lyft-spot/>
3. <https://eng.lyft.com/running-apache-airflow-at-lyft-6e53bb8fccff>

4. <https://expandedramblings.com/index.php/lyft-statistics/>
5. <https://eng.lyft.com/meet-sravanthi-one-of-the-pms-behind-lyfts-new-rider-app-9bc61abe31>