**ASSIGNMENT 2**

**DATA MANAGEMENT AND DATABASE DESIGN**

**PHYSICAL DATA MODEL AND SOCIAL MEDIA**

**INFO6210**

**By:  
TeamASquare  
ANINDITA BAISHYA (NUID: 001387422)  
ABHI PATODI (NUID: 001404833)**

**Part – I**

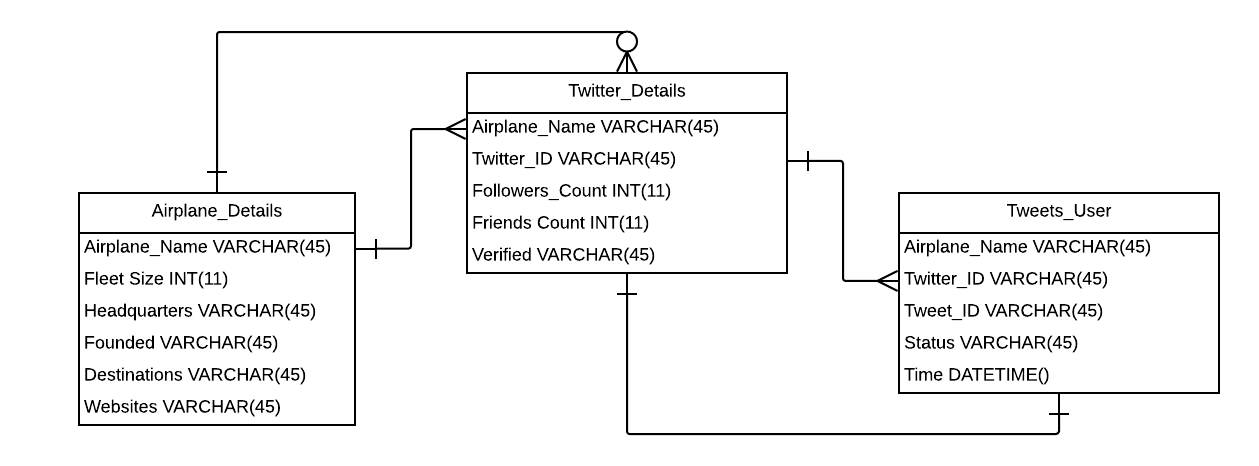
**# Abstract:**

In this assignment, we selected the topic: Airlines as our theme, which is the extension of our first assignment. We are expanding the Airline database in terms of companies, consumers and producers.

* Companies: List of top 10 airlines, their twitter handles and tweets
* Consumers: Reviews, ratings by the consumers on the airlines in Yelp
* Producers: Manufacturer and founder details are show casted

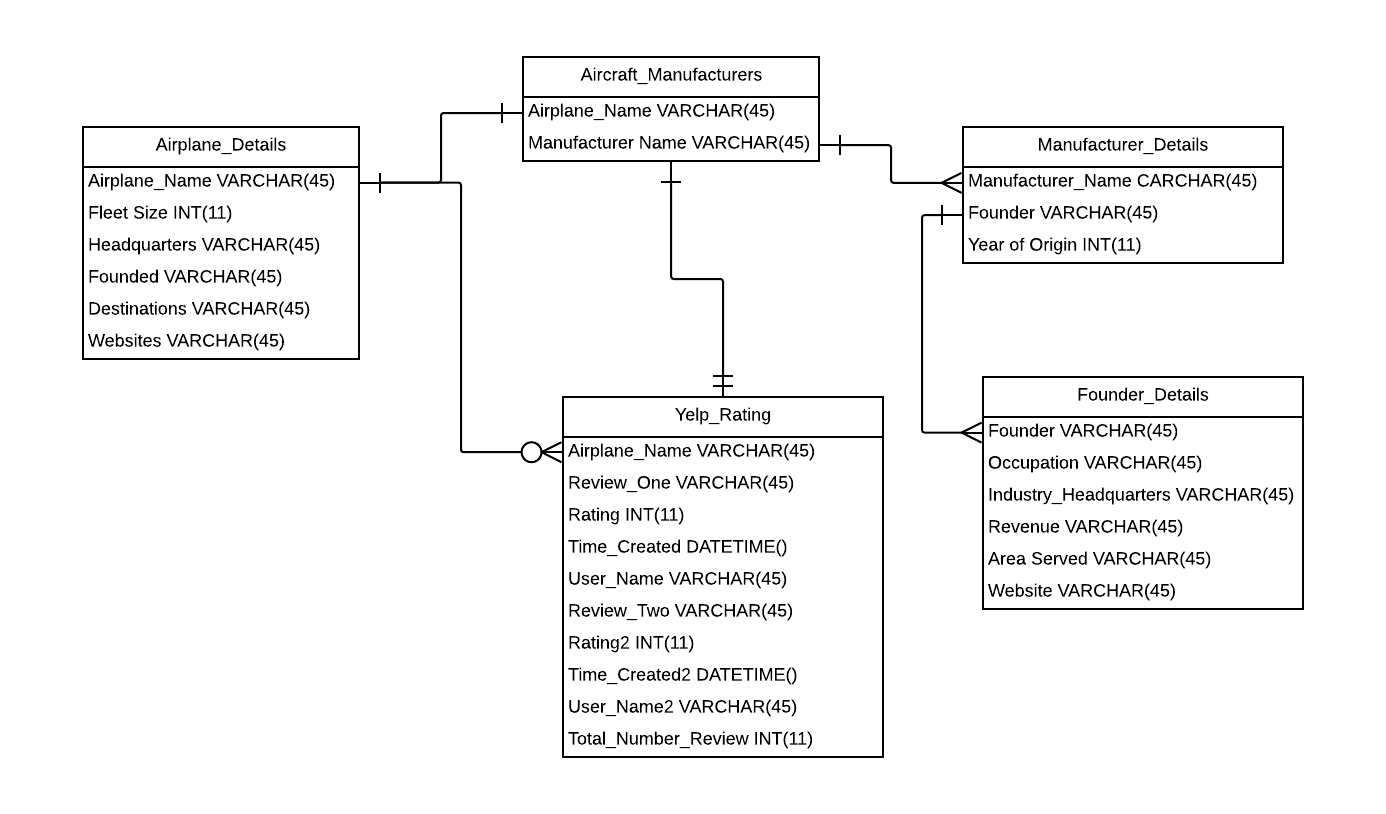
**# Conceptual Models for a tweet/post, Social media handler account of the airline company:**

Below is the conceptual model diagram that represent an airline and their social media user account details along with the tweet related information



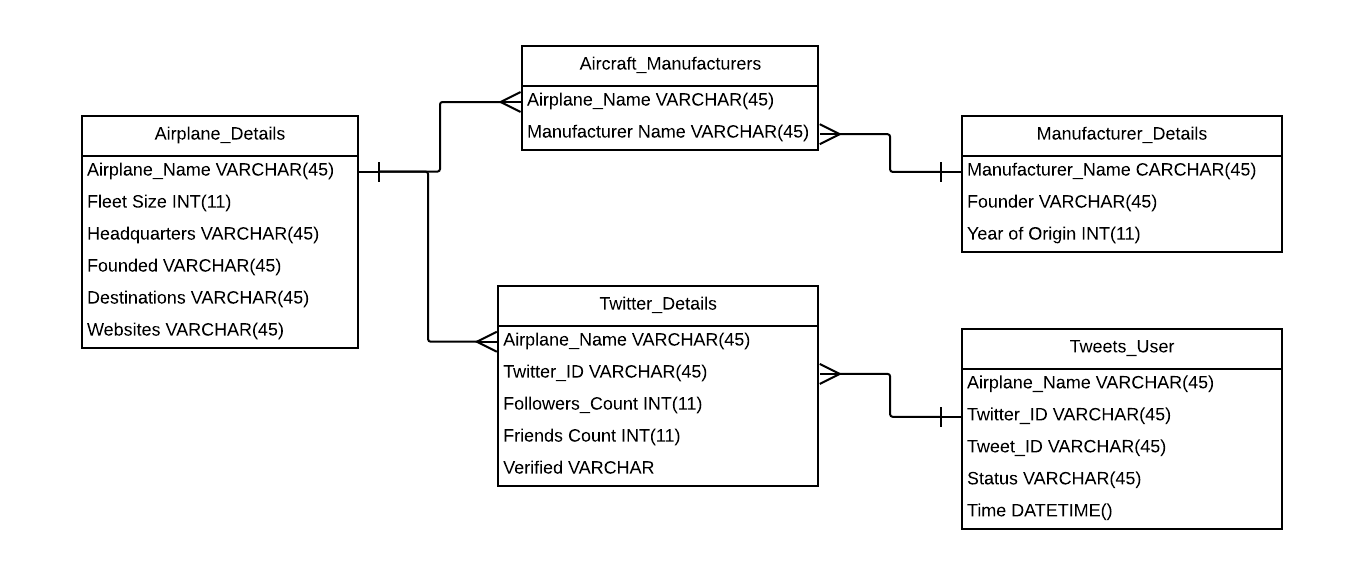
**# Conceptual Model that represents consumers, producers and companies:**

Below is the conceptual model diagram that represent producers i.e. manufacturer details, consumers i.e. yelp users and companies i.e. airline company details.

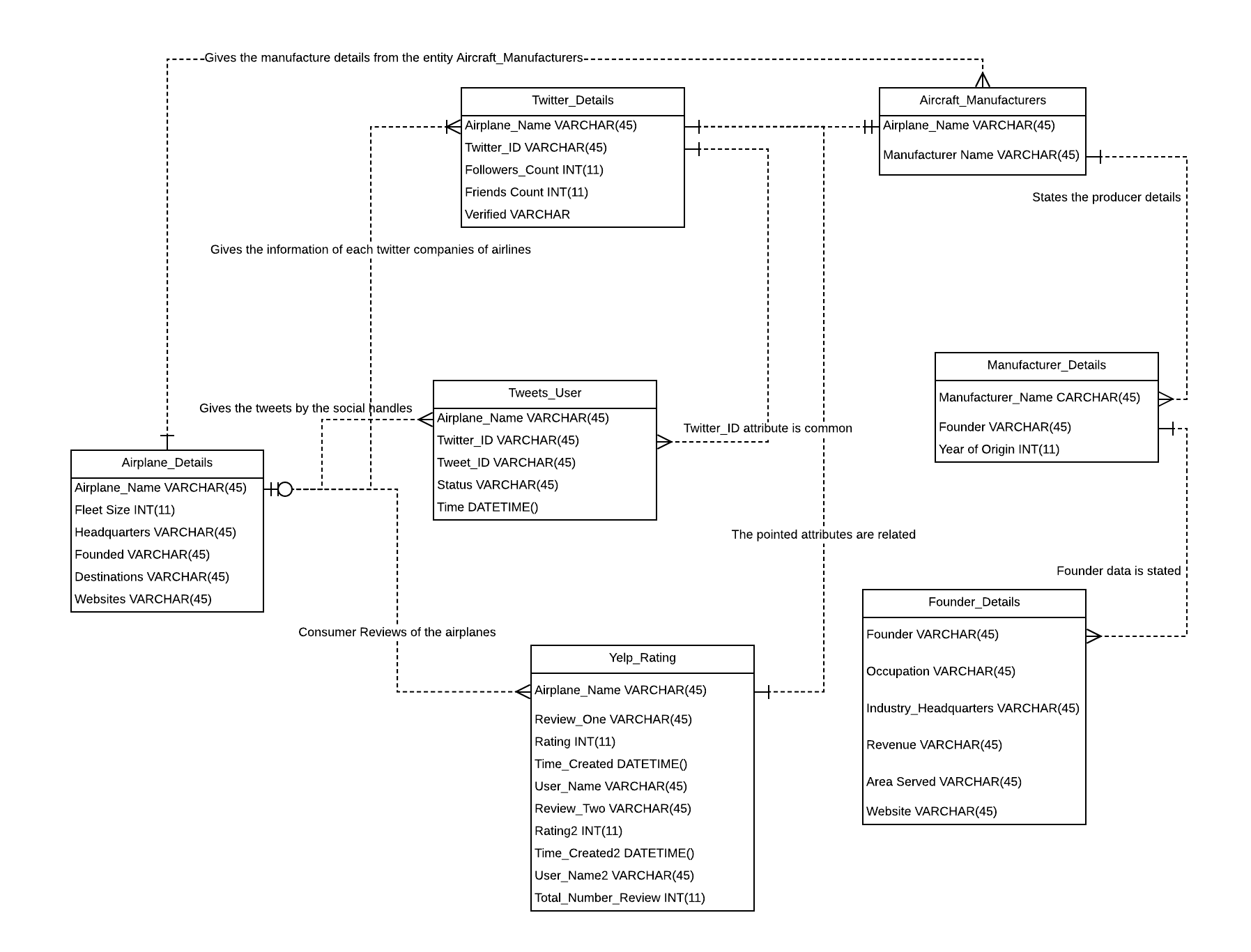


**# Conceptual Models for at least two things specific to the domain:**

Below is the conceptual model diagram that represent two specific things namely Manufacturer details and Twitter details of each airlines which are specific to the airline’s domain.



**# Initial ER Diagram with the entire conceptual model:**



**# Questions to be answered:**

1. ***What are the ranges, data types and format of all the attributes in your entities?***

The ranges, data types and format of all the attributes in our entities are mentioned below:

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | **Varchar(250)** |  | **INT(11)** | **DateTime()** |
| Airplane\_Name | Tweet\_ID | Manufacturer\_Name | Fleet\_Size | Time |
| Headquarters | Status | Founder | Followers\_Count | Time\_Created |
| Founded | Review\_One | Occupation | Friends\_Count | Time\_Created2 |
| Destinations | User\_Name | Revenue | Rating |  |
| Websites | Review\_Two |  | Rating2 |  |
| Twitter\_ID | User\_Name2 |  | Total\_Number\_Review |  |
| Verified | Industry\_Headquarters |  | Fav\_Count |  |
| Website | Area\_Served |  | Year\_of\_Origin |  |

1. ***When should you use an entity versus attribute?***

To understand what information, we are storing inside an entity we use attributes to specify the data types, ranges and the format.

Entities are the things or objects inside a database. And Attributes are the details of the objects or the things called entities.

For eg:

Entity

Aircraft\_Manufacturers

Attributes

1. ***When should you use an entity or relationship, and placement of attributes?***

Relationship between entities is a kind of a mathematical relationship taken from entity sets  
From our ER diagram we can see one such example:

**Airplane\_Name Manufacturer\_Name Founder\_Name**

American Airlines Airbus Franz Josef Strauss

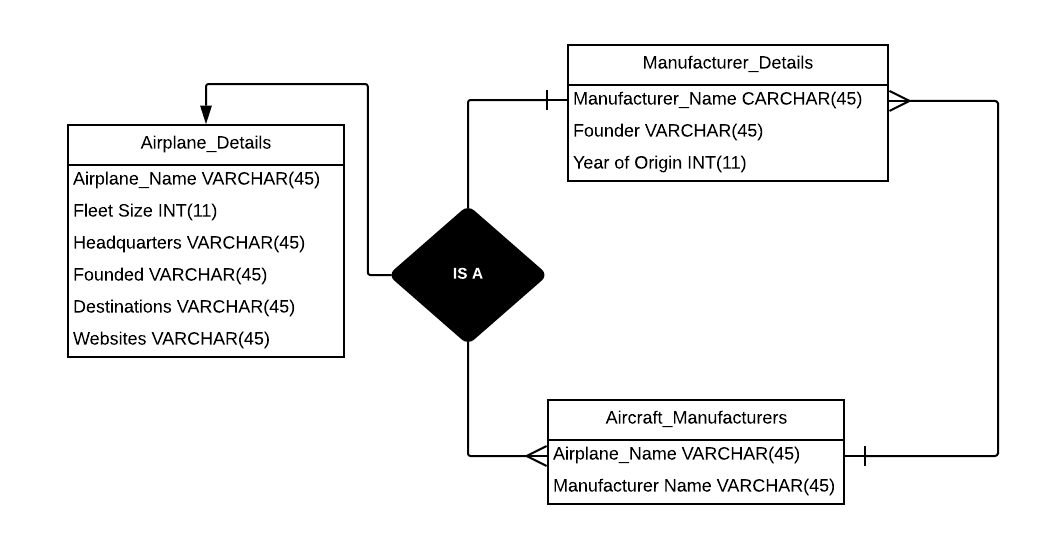
(American Airlines, Franz Josef Strauss) ∈ Manufacturer\_Name (Airbus)

1. ***How did you choose your keys? Which are unique?***

We have chosen the below keys as the unique keys:

* Airplane\_Name: Primary Key (because it is common in all the tables)
* Twitter\_ID: Foreign Key (to link another table for reference and data)
* Manufactuer\_Name: Foreign Key (-do-)
* Founder: Foreign Key (-do-)

1. ***Did you model hierarchies using the “ISA” design element? Why or why not?***

Yes, we used the “ISA” design element while modelling the ER Diagram.  
Because we need to show the relationship between the entities and the best way to break the entities to show their relationship by mapping the tables is by using the ISA design element.One of such example in our ER diagram is: 

1. ***Were there design alternatives? What are their tradeoffs: entity vs. attribute, entity vs. relationship, binary vs. ternary relationship?***

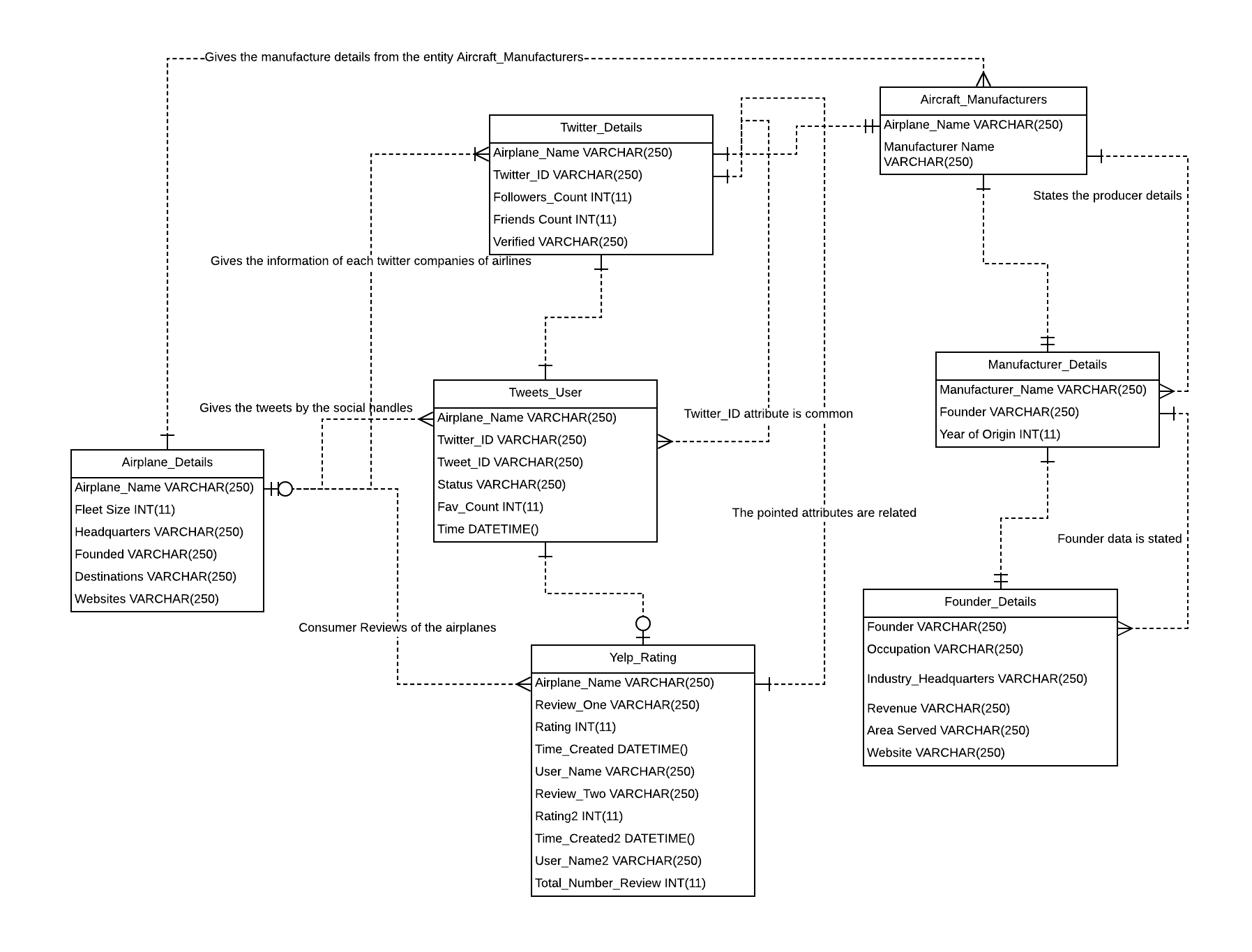
Yes, there are design alternatives in the ER diagram.  
Entity VS. Attribute: Their tradeoff is relationship between the entities are shown via the attributes  
Entity VS. Relationship: Their tradeoff is relationship shown to break the database into components and show relationship between the entities  
Binary VS. Ternary Relationship: Their tradeoff is relationship is shown when two or three entities participate in the design

1. ***Where are you going to find the real-world data to populate your model?***

We have used the below real-world website to gather the data also twitter handles and from Yelp platform.  
<https://bestcompany.com/airlines>

**PART – II**

1. **Updated ER Diagram:**

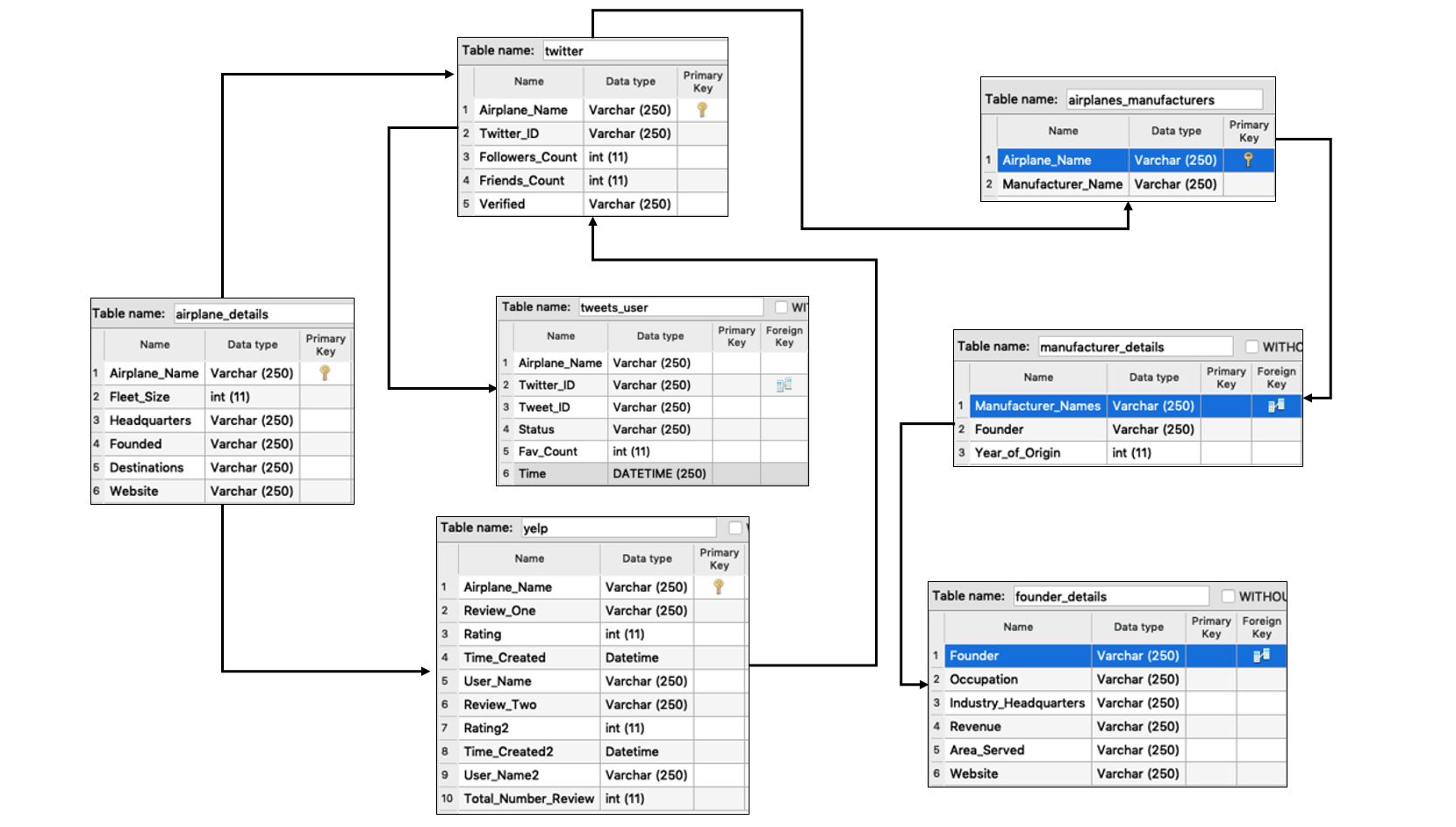


**FEEDBACK:**

* We have interlinked the entity tables that are dependent on each other: (Twitter\_Details, Tweets\_User), (Aircraft\_Manufacturers, Manufacturer\_Details) and (Manufacturer\_Details, Founder\_Details) tables
* In the table Tweets\_User we have inserted a new column Fav\_Count to get the likes in the tweets
* Also, we have increased the data range of the data types VARCHAR from 45 to 250 for capturing tweets of maximum characters

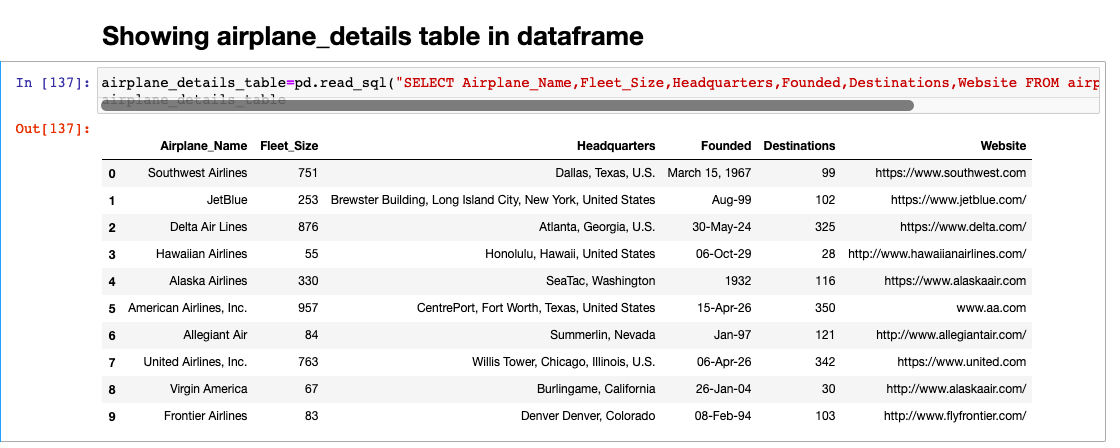
1. **SQL and diagram for the physical model that represents the entire conceptual model.**

SQL tables used to explain the physical model representation of the conceptual model of the Airplanes Database:



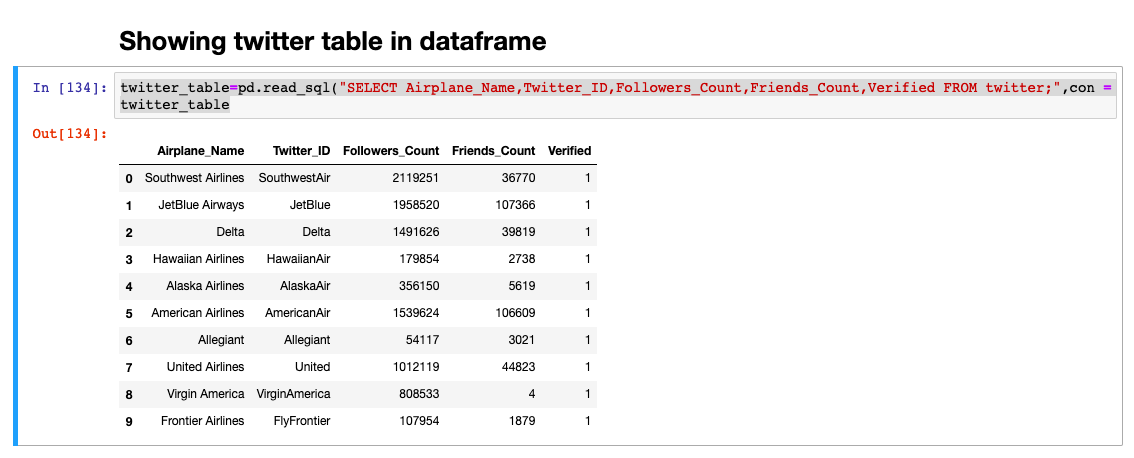
**Table 1: Airplane\_Details**

* SELECT Airplane\_Name, Fleet\_Size, Headquarters, Founded, Destinations, Website FROM Airplane\_Details;



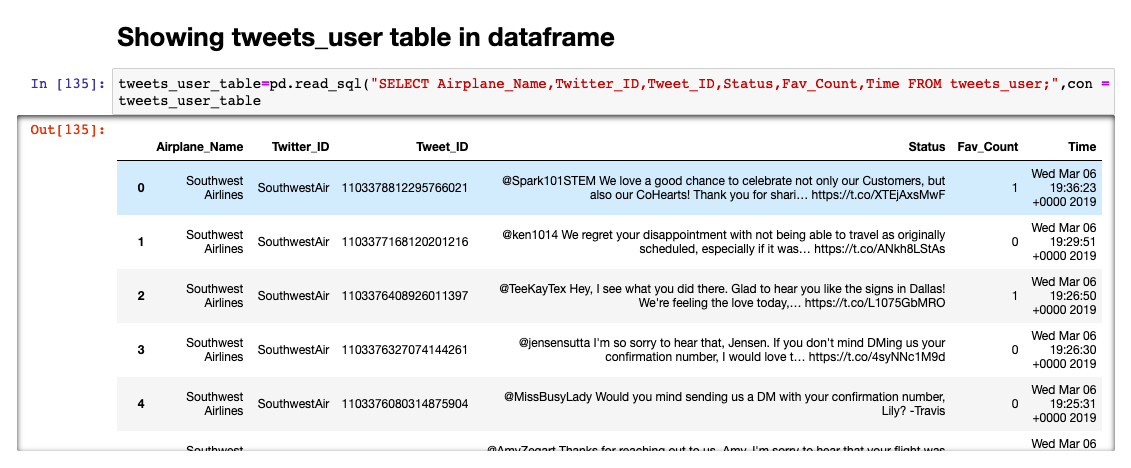
**Table 2: Twitter**

* SELECT Airplane\_Name, Twitter\_ID, Followers\_Count, Friends\_Count, Verified FROM Twitter;



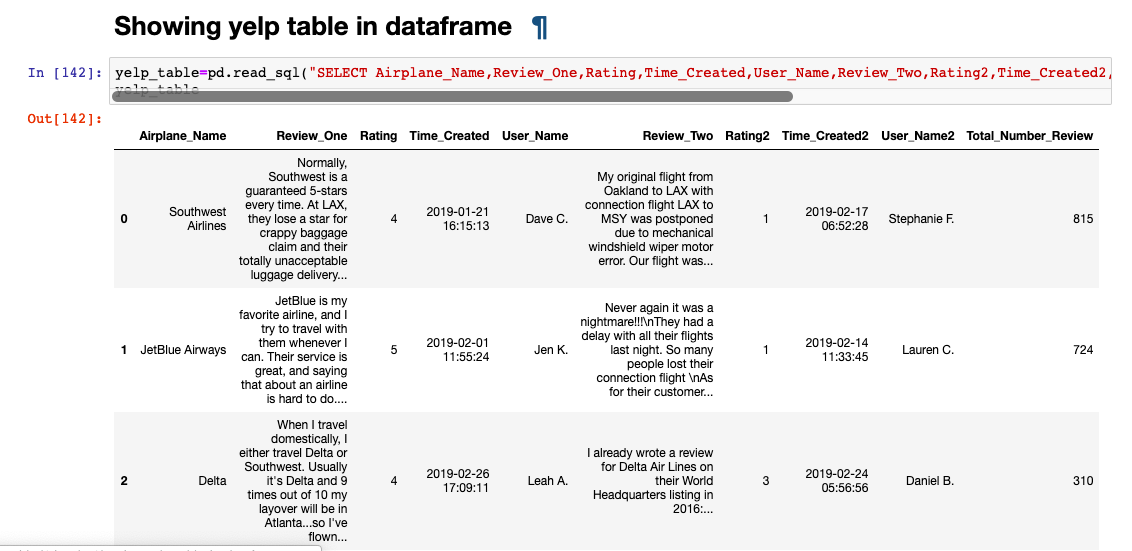
**Table 3: Tweets\_User**

* SELECT Airplane\_Name, Twitter\_ID, Tweet\_ID, Status, Fav\_Count, Time FROM Tweets\_User;



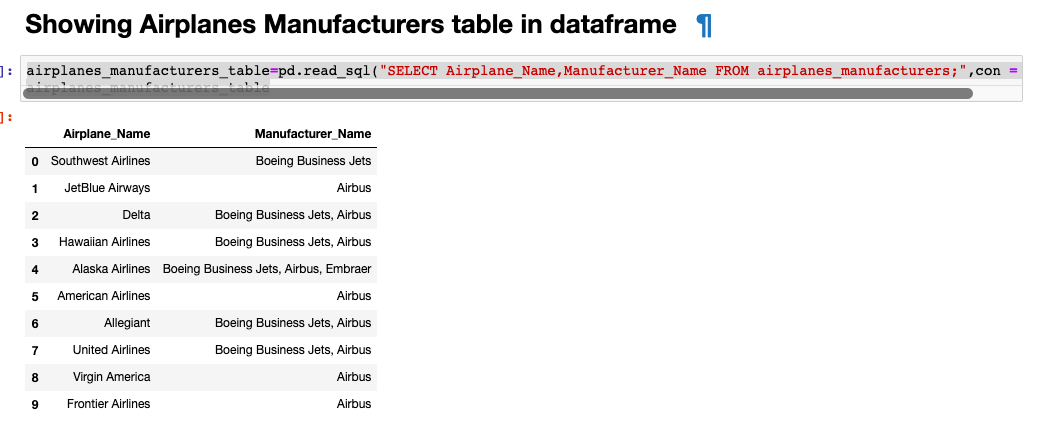
**Table 4: Yelp**

* SELECT Airplane\_Name, Review\_One, Rating, Time\_Created, User\_Name, Review\_Two, Rating2, Time\_Created2, User\_Name2, Total\_Number\_Reviews FROM yelp;



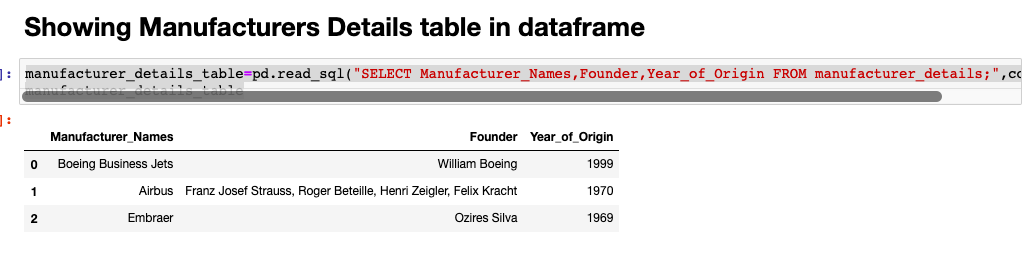
**Table 5: Aircraft\_Manufacturers**

* SELECT Airplane\_Name, Manufacturer\_Name FROM Airplanes\_Manufacturers;



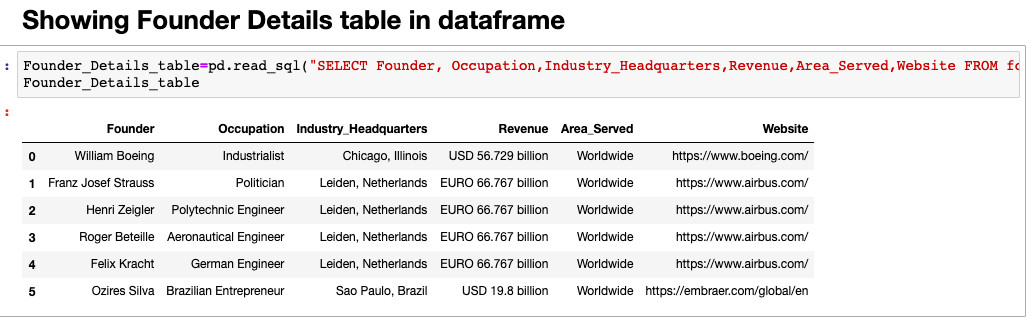
**Table 6: Manufacturer\_Details**

* SELECT Manufacturer\_Names, Founder, Year\_of\_Origin FROM Manufacturer\_Details;



**Table 7: Founder\_Details**

* SELECT Founder, Occupation, Industry\_Headquarters, Revenue, Area\_Served, Website FROM Founder\_Details;



1. **SQL that express the queries you are asked to write.**
2. ***What user posted this (e.g. tweet, facebook post, IG post, etc.)?***

* SELECT Airplane\_Name,Twitter\_ID FROM tweets\_user where Status like '%Can you name this city? If you guessed #Boston, you are correct! As we countdown to the launch of our new non-stop%'



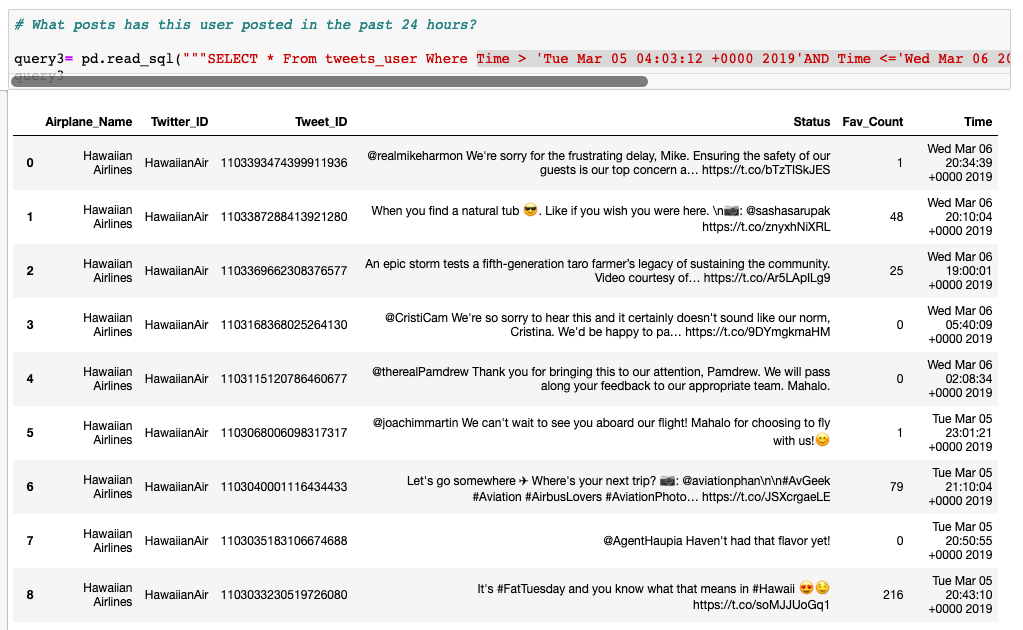
1. ***When did the user post this (e.g. tweet, facebook post, IG post, etc.)??***

* SELECT Airplane\_Name, Twitter\_ID, Time FROM tweets\_user where Status like '%Can you name this city? If you guessed #Boston, you are correct! As we countdown to the launch of our new non-stop%';



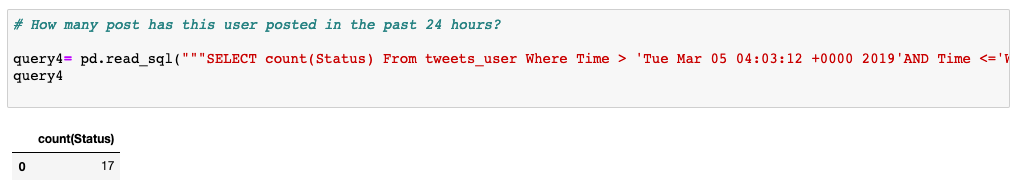
1. ***What posts has this user posted in the past 24 hours?***

* SELECT \* From tweets\_user Where Time > 'Tue Mar 05 04:03:12 +0000 2019'AND Time <='Wed Mar 06 20:34:39 +0000 2019' AND Airplane\_Name='Hawaiian Airlines'""";



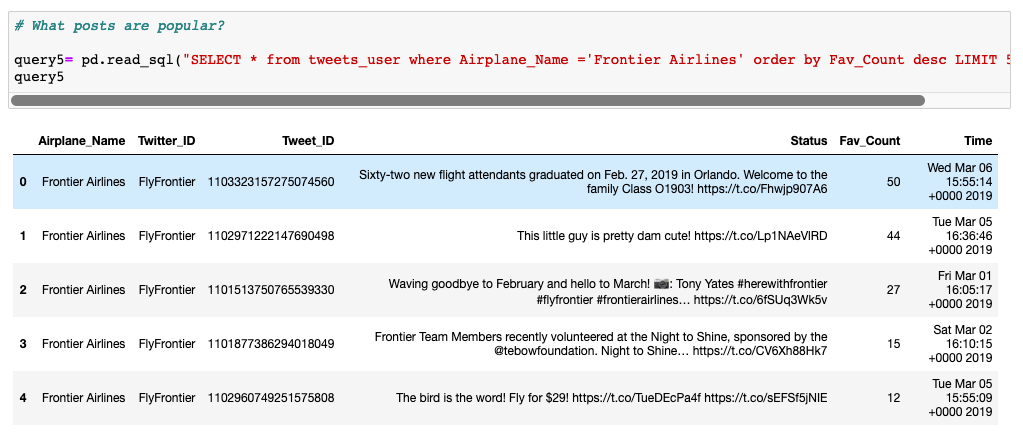
1. ***How many post has this user posted in the past 24 hours?***

* SELECT count(Status) From tweets\_user Where Time > 'Tue Mar 05 04:03:12 +0000 2019'AND Time <='Wed Mar 06 20:34:39 +0000 2019' AND Airplane\_Name='Hawaiian Airlines';



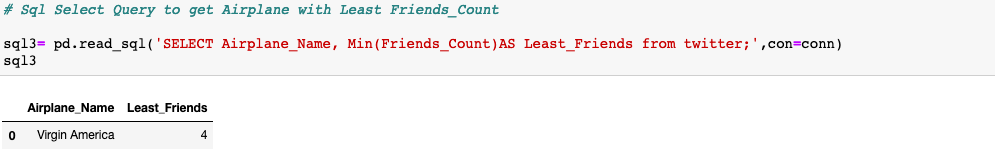
1. ***What posts are popular?***

* SELECT \* from tweets\_user where Airplane\_Name ='Frontier Airlines' order by Fav\_Count desc LIMIT 5;

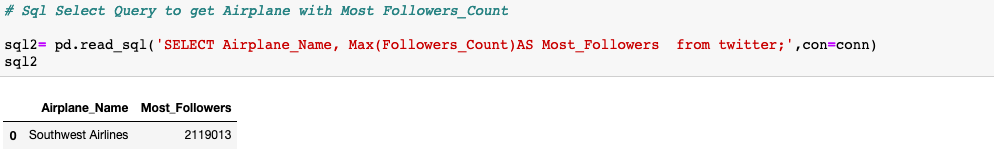


1. **10 USE-CASES particular to our domain:**

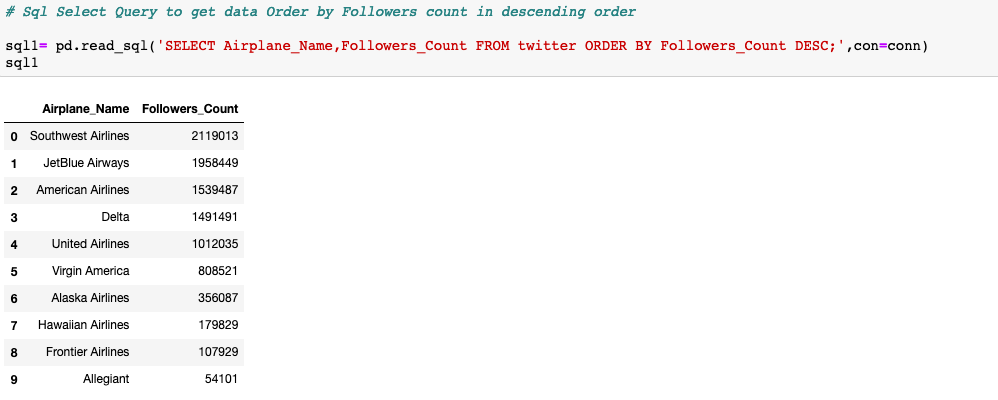
(i). Use-Case query to display Airplane with Least Friends count



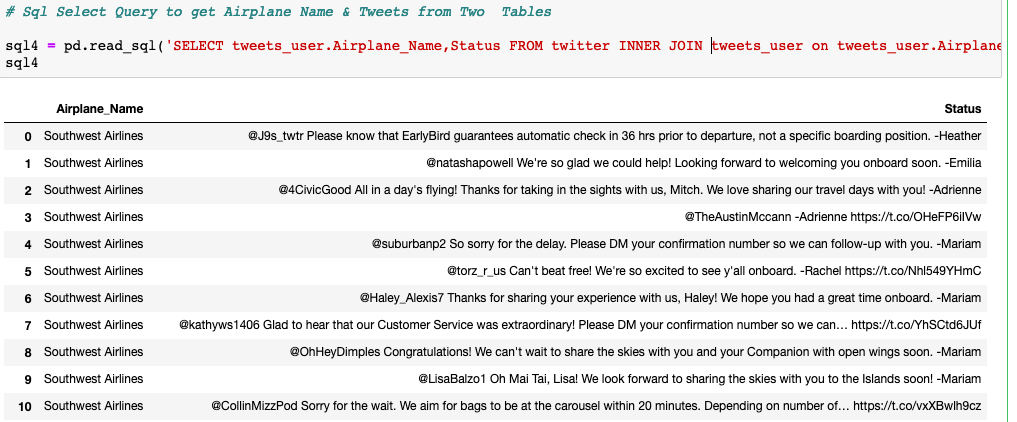
(ii). Use-Case query to display Airplane with Most Followers count



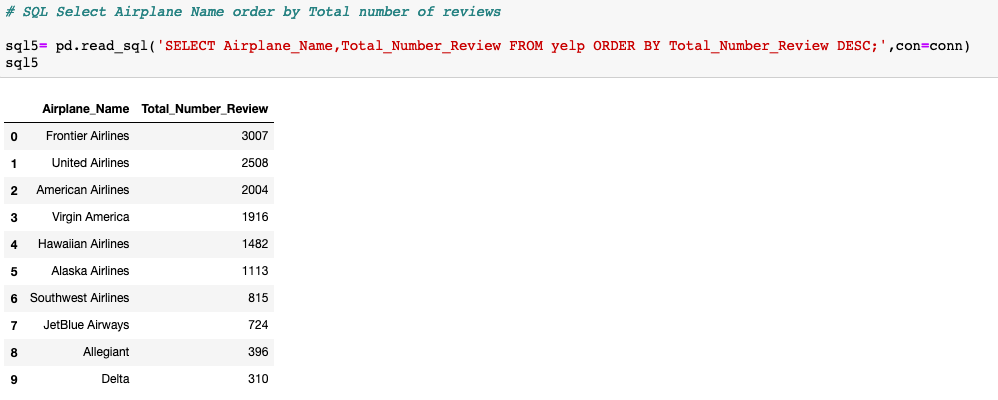
(iii). Use-Case query to display data order by followers count in descending order

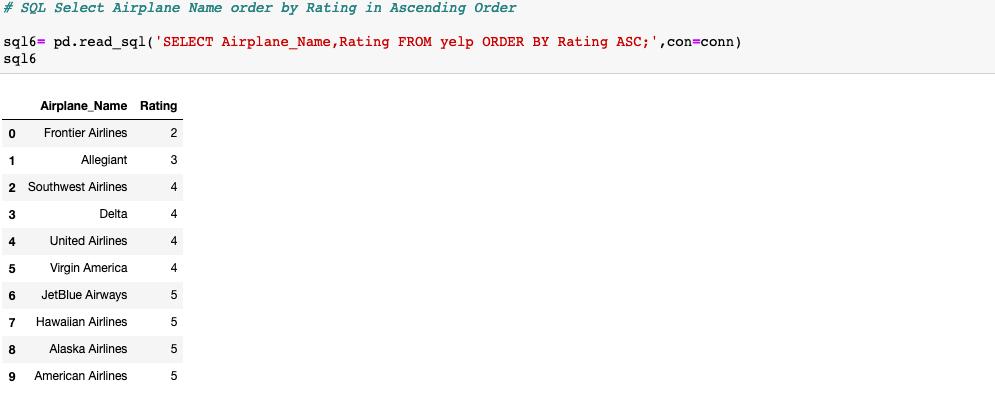


(iv). Use-Case query to display Airplane Name & Tweets from Two Tables

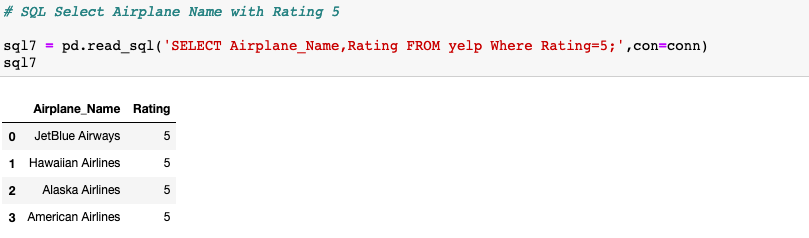


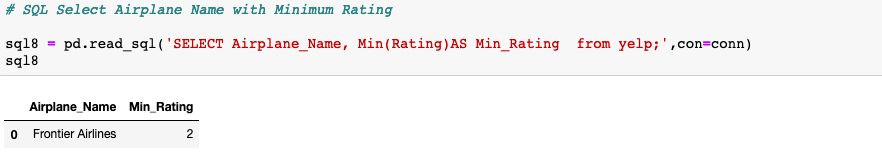
(v). Use-Case query to display Airplane name order by total no of reviews



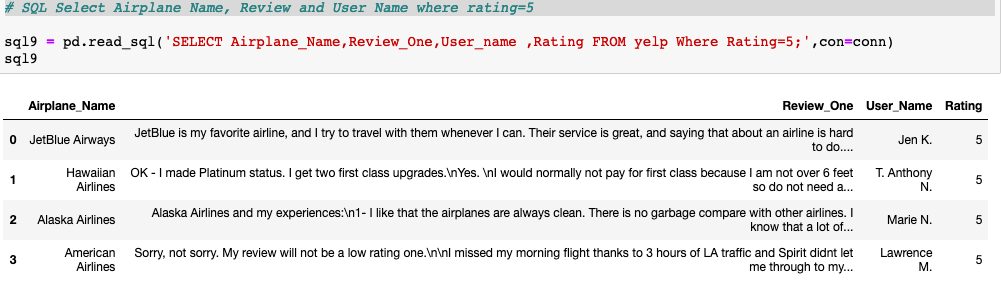
(vi). Use-Case query to display Airplane name order by rating in ascending order  


(vii). Use-Case query to display airplane name with Rating 5

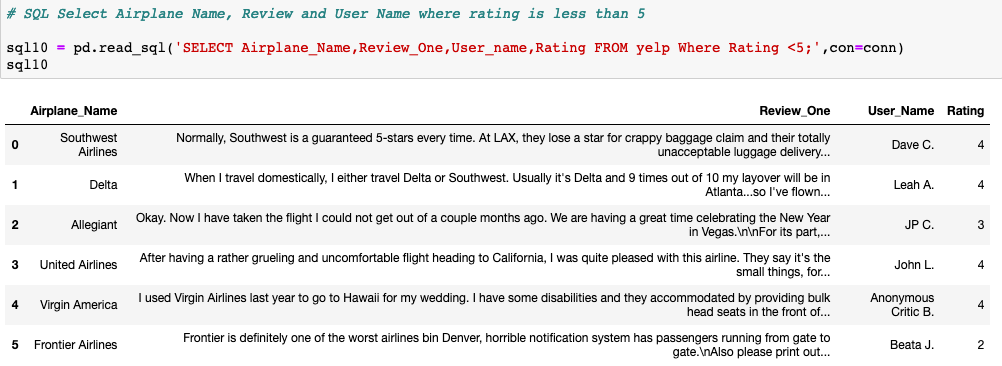


(viii). Use-Case query to display Airplane name with the minimum rating

(ix). Use-Case query to display Airplane Name, Review and User Name who rated 5



(x). Use-Case query to display Airplane Name, Review and User Name who rated less than 5



**Citations and References:**

Each code in this assignment is self-developed and is not copied from any website. Please refer the Jupyter notebook TeamASquare\_Assignment2.ipynb attached along with this document for the code.

References were taken from the below website

<https://stackoverflow.com/>  
<https://www.w3schools.com/sql>  
<https://www.lucidchart.com>  
<https://www.smartdraw.com>

**Contribution:**

Twitter API – Anindita Baishya  
Yelp – Abhi Patodi  
Use Cases – 5 each by Anindita & Abhi   
SQL Queries – Abhi & Anindita  
Conceptual Diagram – Abhi Patodi  
ER Diagram & Documentation – Anindita Baishya

Text License:

This is a human-readable summary of (and not a substitute for) the [license](https://creativecommons.org/licenses/by/3.0/us/legalcode). [Disclaimer](https://creativecommons.org/licenses/by/3.0/us/).

### You are free to:

**Share** — copy and redistribute the material in any medium or format

**Adapt** — remix, transform, and build upon the material for any purpose, even commercially.

The licensor cannot revoke these freedoms as long as you follow the license terms.

### Under the following terms:

**Attribution** — You must give [appropriate credit](https://creativecommons.org/licenses/by/3.0/us/), provide a link to the license, and [indicate if changes were made](https://creativecommons.org/licenses/by/3.0/us/). You may do so in any reasonable manner, but not in any way that suggests the licensor endorses you or your use.

**No additional restrictions** — You may not apply legal terms or [technological measures](https://creativecommons.org/licenses/by/3.0/us/) that legally restrict others from doing anything the license permits.

### Notices:

You do not have to comply with the license for elements of the material in the public domain or where your use is permitted by an applicable [exception or limitation](https://creativecommons.org/licenses/by/3.0/us/).

No warranties are given. The license may not give you all of the permissions necessary for your intended use. For example, other rights such as [publicity, privacy, or moral rights](https://creativecommons.org/licenses/by/3.0/us/) may limit how you use the material.