

**Skills
Network**

([https://skills.network/?](https://skills.network/?utm_medium=Exinfluencer&utm_source=Exinfluencer&utm_content=000026UJ&utm_term=1000SkillsNetwork-Channel-SkillsNetworkCoursesIBMDS0321ENSkillsNetwork26802033-2022-01-01)

[utm_medium=Exinfluencer&utm_source=Exinfluencer&utm_content=000026UJ&utm_term=1000SkillsNetwork-Channel-SkillsNetworkCoursesIBMDS0321ENSkillsNetwork26802033-2022-01-01](https://skills.network/?utm_medium=Exinfluencer&utm_source=Exinfluencer&utm_content=000026UJ&utm_term=1000SkillsNetwork-Channel-SkillsNetworkCoursesIBMDS0321ENSkillsNetwork26802033-2022-01-01))

JSricharan Nallan Chakravarthula 

Space X Falcon 9 First Stage Landing Prediction

Assignment: SQL Notebook for Peer Assignment

Estimated time needed: **60** minutes

Overview of the DataSet

SpaceX has gained worldwide attention for a series of historic milestones.

It is the only private company ever to return a spacecraft from low-earth orbit, which it first accomplished in December 2010. SpaceX advertises Falcon 9 rocket launches on its website with a cost of 62 million dollars whereas other providers cost upward of 165 million dollars each, much of the savings is because Space X can reuse the first stage.

Therefore if we can determine if the first stage will land, we can determine the cost of a launch.

This information can be used if an alternate company wants to bid against SpaceX for a rocket launch.

This dataset includes a record for each payload carried during a SpaceX mission into outer space.

Download the datasets

This assignment requires you to load the spacex dataset.

In many cases the dataset to be analyzed is available as a .CSV (comma separated values) file, perhaps on the internet. Click on the link below to download and save the dataset (.CSV file):

[Spacex DataSet \(https://cf-courses-data.s3.us.cloud-object-storage.appdomain.cloud/IBM-DS0321EN-SkillsNetwork/labs/module_2/data/Spacex.csv?utm_medium=Exinfluencer&utm_source=Exinfluencer&utm_content=000026UJ&utm_term=1000SkillsNetwork-Channel-SkillsNetworkCoursesIBMD�0321ENSkillsNetwork26802033-2022-01-01\)](https://cf-courses-data.s3.us.cloud-object-storage.appdomain.cloud/IBM-DS0321EN-SkillsNetwork/labs/module_2/data/Spacex.csv?utm_medium=Exinfluencer&utm_source=Exinfluencer&utm_content=000026UJ&utm_term=1000SkillsNetwork-Channel-SkillsNetworkCoursesIBMD�0321ENSkillsNetwork26802033-2022-01-01)

Store the dataset in database table

it is highly recommended to manually load the table using the database console LOAD tool in DB2.

LOAD DATA

Source Target Define Finalize

You are loading the file *Spacex.csv*

Select a load target

Schema ⊕ New Schema Find a schema

Table ⊕ New Table Find a table in QWP24135

Refresh

Schema	Table
AUDIT	ANNUAL_CROP_DATA
DB2INST1	BOARD
ERRORSCHEMA <i>Sample</i>	BOOKSHOP
IDAX	BOOKSHOP_AUTHORDetails
QWP24135	CAR_SALES
SQL15777	CAR_SALES_DATA

Create a new Table

SPACEXTBL

Create

Back Next

Now open the Db2 console, open the LOAD tool, Select / Drag the .CSV file for the dataset, Next create a New Table, and then follow the steps on-screen instructions to load the data. Name the new table as follows:

SPACEXDATASET

Follow these steps while using old DB2 UI which is having Open Console Screen

Note:While loading Spacex dataset, ensure that detect datatypes is disabled. Later click on the pencil icon(edit option).

1. Change the Date Format by manually typing DD-MM-YYYY and timestamp format as DD-MM-YYYY HH:MM:SS
2. Change the PAYLOAD_MASS__KG_ datatype to INTEGER.

LOAD DATA

Source Target Define Finalize

You are loading the file **Spacex.csv** into **QWP24135.SPACEXTBL**

Code page (character encoding): **1208 (UTF-8)** Separator: **,** Header in first row: ☒ Time & date format: **DD-MM-YYYY HH-MM-SS** Detect data types: ☐

Date format: **DD-MM-YYYY** Time format: **HH-MM-SS** Timestamp format: **DD-MM-YYYY HH-MM-SS**

Changes to be considered when having DB2 instance with the new UI having Go to UI screen

- Refer to this instruction in this [link \(https://cf-courses-data.s3.us.cloud-object-storage.appdomain.cloud/IBMDeveloperSkillsNetwork-DB0201EN-SkillsNetwork/labs/Labs_Coursera_V5/labs/Lab%20-%20Sign%20up%20for%20IBM%20Cloud%20-%20Create%20Db2%20service%20instance%20-%20Get%20started%20with%20the%20Db2%20console/instructional-labs.md.html?utm_medium=Exinfluencer&utm_source=Exinfluencer&utm_content=000026UJ&utm_term=SkillsNetwork-Channel-SkillsNetworkCoursesIBMD50321ENSkillsNetwork26802033-2022-01-01\)](https://cf-courses-data.s3.us.cloud-object-storage.appdomain.cloud/IBMDeveloperSkillsNetwork-DB0201EN-SkillsNetwork/labs/Labs_Coursera_V5/labs/Lab%20-%20Sign%20up%20for%20IBM%20Cloud%20-%20Create%20Db2%20service%20instance%20-%20Get%20started%20with%20the%20Db2%20console/instructional-labs.md.html?utm_medium=Exinfluencer&utm_source=Exinfluencer&utm_content=000026UJ&utm_term=SkillsNetwork-Channel-SkillsNetworkCoursesIBMD50321ENSkillsNetwork26802033-2022-01-01) for viewing the new Go to UI screen.
- Later click on **Data link(below SQL)** in the Go to UI screen and click on **Load Data** tab.
- Later browse for the downloaded spacex file.

IBM Db2 on Cloud

Load Data Load History Tables Views Indexes Aliases MQTs Sequences Application objects

Source Target Define Finalize

You are loading the file

File selection

My Computer
A single delimited text file (CSV) without header row.

Amazon S3

Cloud Object Storage

Drag a file here or [browse files](#)

- Once done select the schema and load the file.

SQL

Source Target Define Finalize

You are loading the file **Spacex (2).csv** into **SRW76180.SPACEXTBL**

Code page (character encoding): **1208 (UTF-8)** Separator: **,** Header in first row: ☒ Time & date format: **DD-MM-YYYY HH-MM-SS**

Date format: **DD-MM-YYYY** Time format: **HH-MM-SS** Timestamp format: **DD-MM-YYYY HH-MM-SS**

	DATE DATE	TIME UTC TIME	BOOSTER_VERSION VARCHAR	LAUNCH_SITE VARCHAR	PAYLOAD VARCHAR	PAYLOAD SMALLINT
1	04-06-2010	18:45:00	F9 v1.0 B0003	CCAFS LC-40	Dragon Spacecraft Qualification Unit	0
2	08-12-2010	15:43:00	F9 v1.0 B0004	CCAFS LC-40	Dragon demo flight C1, two CubeSats, barrel of Brouere cheese	0
3	22-05-2012	07:44:00	F9 v1.0 B0005	CCAFS LC-40	Dragon demo flight C2	525
4	08-10-2012	00:35:00	F9 v1.0 B0006	CCAFS LC-40	SpaceX CRS-1	500
5	01-03-2013	15:10:00	F9 v1.0 B0007	CCAFS LC-40	SpaceX CRS-2	677
6	29-09-2013	16:00:00	F9 v1.1 B1003	VAFB SLC-4E	CASSIOPE	500
7	03-12-2013	22:41:00	F9 v1.1	CCAFS LC-40	SES-8	3170
8	06-01-2014	22:06:00	F9 v1.1	CCAFS LC-40	ThaiCom 6	3325

Back Next

Objectives

- Understand the SpaceX DataSet
- Load the dataset into a table in a Db2 database
- Execute SQL queries to answer assignment questions

Import Libraries

```
In [1]: ► !pip install sqlalchemy==1.3.9
!pip install -q pandas==1.1.5
print("Installation complete.")
```

Requirement already satisfied: sqlalchemy==1.3.9 in /home/jupyterlab/conda/envs/python/lib/python3.7/site-packages (1.3.9)
Installation complete.

Start Here

Connect to the database

Let us first load the SQL extension and establish a connection with the database

```
In [2]: ► %load_ext sql
print("Done.")
```

Done.

```
In [ ]: ► # Connection string with credentials to IBM Db2
# %sql ibm_db_sa://my-username:my-password@hostname:port/BLUDB?security=SSL
%sql ibm_db_sa://my-username:my-password@hostname:port/bludb?security=SSL
```

```
In [4]: ► # Number of columns
%sql SELECT colcount AS "Number of Columns" FROM syscat.tables WHERE tabna

* ibm_db_sa://zqs31076:***@9938aec0-8105-433e-8bf9-0fbb7e483086.c1ogj3sd
0tgtu0lqde00.databases.appdomain.cloud:32459/bludb
Done.
```

Out[4]: Number of Columns

```
In [5]: ► # Number of Rows
%sql SELECT count(*) AS "Number of Rows" FROM SPACEXDATASET;

* ibm_db_sa://zqs31076:***@9938aec0-8105-433e-8bf9-0fbb7e483086.c1ogj3sd
0tgtu0lqde00.databases.appdomain.cloud:32459/bludb
Done.
```

Out[5]: **Number of Rows**

101

```
In [6]: ► # Column names and info
%sql SELECT colno, colname, typename, length FROM syscat.columns WHERE tab

* ibm_db_sa://zqs31076:***@9938aec0-8105-433e-8bf9-0fbb7e483086.c1ogj3sd
0tgtu0lqde00.databases.appdomain.cloud:32459/bludb
Done.
```

Out[6]:

	colno	colname	typename	length
	0	DATE	DATE	4
	1	TIME__UTC__	TIME	3
	2	BOOSTER_VERSION	VARCHAR	14
	3	LAUNCH_SITE	VARCHAR	12
	4	PAYLOAD	VARCHAR	61
	5	PAYLOAD_MASS__KG__	SMALLINT	2
	6	ORBIT	VARCHAR	11
	7	CUSTOMER	VARCHAR	57
	8	MISSION_OUTCOME	VARCHAR	32
	9	LANDING__OUTCOME	VARCHAR	22

```
In [7]: # First 4 rows
%sql SELECT * FROM SPACEXDATASET LIMIT 4;

* ibm_db_sa://zqs31076:***@9938aec0-8105-433e-8bf9-0fbb7e483086.c1ogj3sd
0tgtu0lqde00.databases.appdomain.cloud:32459/bludb
Done.
```

Out[7]:

DATE	time__utc__	booster_version	launch_site	payload	payload_mass__kg__	orbit	cus
2010-06-04	18:45:00	F9 v1.0 B0003	CCAFS LC-40	Dragon Spacecraft Qualification Unit	0	LEO	S
2010-12-08	15:43:00	F9 v1.0 B0004	CCAFS LC-40	Dragon demo flight C1, two CubeSats, barrel of Brouere cheese	0	LEO (ISS)	(
2012-05-22	07:44:00	F9 v1.0 B0005	CCAFS LC-40	Dragon demo flight C2	525	LEO (ISS)	(
2012-10-08	00:35:00	F9 v1.0 B0006	CCAFS LC-40	SpaceX CRS-1	500	LEO (ISS)	

Solve the Assignment Tasks

Tasks

Now write and execute SQL queries to solve the assignment tasks.

Note: If the column names are in mixed case enclose it in double quotes For Example "Landing_Outcome"

Task 1

Task 1

Display the names of the unique launch sites in the space mission

```
In [8]: # Task 1 ANSWER
%sql SELECT DISTINCT LAUNCH_SITE FROM SPACEXDATASET;

* ibm_db_sa://zqs31076:***@9938aec0-8105-433e-8bf9-0fbb7e483086.c1ogj3sd
0tgtu0lqde00.databases.appdomain.cloud:32459/bludb
Done.
```

```
Out[8]: launch_site
        CCAFS LC-40
        CCAFS SLC-40
        KSC LC-39A
        VAFB SLC-4E
```

Task 2

Task 2

Display 5 records where launch sites begin with the string 'CCA'

```
In [9]: # Task 2 ANSWER
%sql SELECT * FROM SPACEXDATASET WHERE launch_site LIKE 'CCA%' LIMIT 5;

* ibm_db_sa://zqs31076:***@9938aec0-8105-433e-8bf9-0fbb7e483086.c1ogj3sd
0tgtu0lqde00.databases.appdomain.cloud:32459/bludb
Done.
```

```
Out[9]: DATE    time_utc_  booster_version  launch_site  payload  payload_mass_kg_  orbit  cus

2010-06-04  18:45:00  F9 v1.0 B0003  CCAFS LC-40  Dragon Spacecraft Qualification Unit  0  LEO  S

2010-12-08  15:43:00  F9 v1.0 B0004  CCAFS LC-40  Dragon demo flight C1, two CubeSats, barrel of Brouere cheese  0  LEO (ISS)  (

2012-05-22  07:44:00  F9 v1.0 B0005  CCAFS LC-40  Dragon demo flight C2  525  LEO (ISS)  (

2012-10-08  00:35:00  F9 v1.0 B0006  CCAFS LC-40  SpaceX CRS-1  500  LEO (ISS)

2013-03-01  15:10:00  F9 v1.0 B0007  CCAFS LC-40  SpaceX CRS-2  677  LEO (ISS)
```

Task 3

Task 3

Display the total payload mass carried by boosters launched by NASA (CRS)

```
In [10]: ▶ # Task 3 ANSWER
%sql SELECT sum(payload_mass__kg_) AS "Total Payload Mass (kg)" FROM SPACE_
* ibm_db_sa://zqs31076:***@9938aec0-8105-433e-8bf9-0fbb7e483086.c1ogj3sd
0tgtu0lqde00.databases.appdomain.cloud:32459/bludb
Done.
```

Out[10]: **Total Payload Mass (kg)**

48213

Task 4

Task 4

Display average payload mass carried by booster version F9 v1.1

```
In [11]: ▶ # Task 4 ANSWER
%sql SELECT sum(payload_mass__kg_) / count(payload_mass__kg_) AS "Average"
* ibm_db_sa://zqs31076:***@9938aec0-8105-433e-8bf9-0fbb7e483086.c1ogj3sd
0tgtu0lqde00.databases.appdomain.cloud:32459/bludb
Done.
```

Out[11]: **Average Payload Mass (kg)**

2928

Task 5

Task 5

List the date when the first successful landing outcome in ground pad was achieved.

Hint: Use min function




```
In [12]: # Task 5 ANSWER
%sql SELECT min(DATE) AS "First Successful Landing Outcome Date" FROM SPACEXDATASET

* ibm_db_sa://zqs31076:***@9938aec0-8105-433e-8bf9-0fbb7e483086.c1ogj3sd0tgtu0lqde00.databases.appdomain.cloud:32459/bludb
Done.
```

```
Out[12]: First Successful Landing Outcome Date
2015-12-22
```

```
In [13]: # This is the full record of the first successful landing outcome in ground
%sql SELECT * FROM SPACEXDATASET WHERE DATE = '2015-12-22';

* ibm_db_sa://zqs31076:***@9938aec0-8105-433e-8bf9-0fbb7e483086.c1ogj3sd0tgtu0lqde00.databases.appdomain.cloud:32459/bludb
Done.
```

```
Out[13]: DATE time__utc_ booster_version launch_site payload payload_mass_kg_ orbit custome
2015-12-22 01:29:00 F9 FT B1019 CCAFS LC-40 OG2 Mission 2 11 Orbcomm-OG2 satellites
2034 LEO Orbc
```

Task 6

Task 6

List the names of the boosters which have success in drone ship and have payload mass greater than 4000 but less than 6000

```
In [14]: # Task 6 ANSWER
%sql SELECT DISTINCT booster_version FROM SPACEXDATASET WHERE landing_outcome = 'Drone Ship'

* ibm_db_sa://zqs31076:***@9938aec0-8105-433e-8bf9-0fbb7e483086.c1ogj3sd0tgtu0lqde00.databases.appdomain.cloud:32459/bludb
Done.
```

```
Out[14]: booster_version
F9 FT B1021.2
F9 FT B1031.2
F9 FT B1022
F9 FT B1026
```



```
In [18]: ▶ # Maximum Payload Mass (kg)
%sql SELECT max(payload_mass__kg_) AS "Max Payload Mass (kg)" FROM SPACEXD
* ibm_db_sa://zqs31076:***@9938aec0-8105-433e-8bf9-0fbb7e483086.c1ogj3sd
0tgtu0lqde00.databases.appdomain.cloud:32459/bludb
Done.
```

Out[18]: **Max Payload Mass (kg)**

15600

```
In [19]: ▶ # Task 8 ANSWER
%sql SELECT booster_version, payload_mass__kg_ FROM SPACEXDATASET WHERE pa
◀────────────────────────────────────────────────────────────────────────▶
* ibm_db_sa://zqs31076:***@9938aec0-8105-433e-8bf9-0fbb7e483086.c1ogj3sd
0tgtu0lqde00.databases.appdomain.cloud:32459/bludb
Done.
```


Out[19]: **booster_version payload_mass__kg_**

F9 B5 B1048.4	15600
F9 B5 B1049.4	15600
F9 B5 B1051.3	15600
F9 B5 B1056.4	15600
F9 B5 B1048.5	15600
F9 B5 B1051.4	15600
F9 B5 B1049.5	15600
F9 B5 B1060.2	15600
F9 B5 B1058.3	15600
F9 B5 B1051.6	15600
F9 B5 B1060.3	15600
F9 B5 B1049.7	15600

Task 9

Task 9

List the records which will display the month names, failure landing_outcomes in drone ship, booster versions, launch_site for the months in year 2015.


In [20]:  *# Number of Unique Launch Dates*

```
%sql SELECT count(DISTINCT Date) AS "Number of Unique Launch Dates" FROM S
```

```
* ibm_db_sa://zqs31076:***@9938aec0-8105-433e-8bf9-0fbb7e483086.c1ogj3sd
0tgtu0lqde00.databases.appdomain.cloud:32459/bludb
Done.
```

Out[20]: **Number of Unique Launch Dates**

101


In [21]:  *# Selecting Year, Month, Day, and Day of week from 'Date' information.*

```
%sql SELECT DATE, YEAR(Date) AS "Year", MONTH(Date) AS "Month", DAY(Date)
```

```
* ibm_db_sa://zqs31076:***@9938aec0-8105-433e-8bf9-0fbb7e483086.c1ogj3sd
0tgtu0lqde00.databases.appdomain.cloud:32459/bludb
Done.
```

Out[21]:

DATE	Year	Month	Day	day_of_week
2010-06-04	2010	6	4	6
2010-12-08	2010	12	8	4
2012-05-22	2012	5	22	3
2012-10-08	2012	10	8	2
2013-03-01	2013	3	1	6

In [22]:  *# Selecting Year, Month, Day, and Day of week from 'Date' information.*

```
%sql SELECT DATE, YEAR(Date) AS "Year", MONTHNAME(Date) AS "Month", DAY(Date)
```

```
* ibm_db_sa://zqs31076:***@9938aec0-8105-433e-8bf9-0fbb7e483086.c1ogj3sd
0tgtu0lqde00.databases.appdomain.cloud:32459/bludb
Done.
```

Out[22]:

DATE	Year	Month	Day	day_of_week	day_of_week_1
2010-06-04	2010	June	4	6	Friday
2010-12-08	2010	December	8	4	Wednesday
2012-05-22	2012	May	22	3	Tuesday
2012-10-08	2012	October	8	2	Monday
2013-03-01	2013	March	1	6	Friday

```
In [23]: ► # All Landing outcomes from 2015
%sql SELECT MONTHNAME(DATE) AS "Month", landing__outcome, booster_version,
* ibm_db_sa://zqs31076:***@9938aec0-8105-433e-8bf9-0fbb7e483086.c1ogj3sd
0tgtu0lqde00.databases.appdomain.cloud:32459/bludb
Done.
```

```
Out[23]:
```

Month	landing__outcome	booster_version	launch_site
January	Failure (drone ship)	F9 v1.1 B1012	CCAFS LC-40
February	Controlled (ocean)	F9 v1.1 B1013	CCAFS LC-40
March	No attempt	F9 v1.1 B1014	CCAFS LC-40
April	Failure (drone ship)	F9 v1.1 B1015	CCAFS LC-40
April	No attempt	F9 v1.1 B1016	CCAFS LC-40
June	Precluded (drone ship)	F9 v1.1 B1018	CCAFS LC-40
December	Success (ground pad)	F9 FT B1019	CCAFS LC-40

```
In [24]: ► # Task 9 ANSWER
%sql SELECT MONTHNAME(DATE) AS "Month", landing__outcome, booster_version,
* ibm_db_sa://zqs31076:***@9938aec0-8105-433e-8bf9-0fbb7e483086.c1ogj3sd
0tgtu0lqde00.databases.appdomain.cloud:32459/bludb
Done.
```

```
Out[24]:
```

Month	landing__outcome	booster_version	launch_site
January	Failure (drone ship)	F9 v1.1 B1012	CCAFS LC-40
April	Failure (drone ship)	F9 v1.1 B1015	CCAFS LC-40

Task 10

Task 10

Rank the count of successful landing_outcomes between the date 04-06-2010 and 20-03-2017 in descending order.

```
In [25]: ► # Task 10 ANSWER
# Rank the counts (of SUCCESSFUL landing_outcomes between the date 04-06-2010 and 04-06-2011)
%sql SELECT landing_outcome, count(landing_outcome) AS "Count" FROM SPACESHIP_LANDINGS
* ibm_db_sa://zqs31076:***@9938aec0-8105-433e-8bf9-0fbb7e483086.c1ogj3sd0tgtu0lqde00.databases.appdomain.cloud:32459/bludb
Done.
```

```
Out[25]:
```

landing__outcome	Count
Success (drone ship)	5
Success (ground pad)	3

```
In [26]: ► # PowerPoint Slide Answer (Includes ALL landing outcomes, not just the successful ones)
# Rank the counts (of ALL landing_outcomes between the date 04-06-2010 and 04-06-2011)
%sql SELECT landing_outcome, count(landing_outcome) AS "Count" FROM SPACESHIP_LANDINGS
* ibm_db_sa://zqs31076:***@9938aec0-8105-433e-8bf9-0fbb7e483086.c1ogj3sd0tgtu0lqde00.databases.appdomain.cloud:32459/bludb
Done.
```

```
Out[26]:
```

landing__outcome	Count
No attempt	10
Failure (drone ship)	5
Success (drone ship)	5
Controlled (ocean)	3
Success (ground pad)	3
Failure (parachute)	2
Uncontrolled (ocean)	2
Precluded (drone ship)	1

End Here