**Analysis of Amazon Books**

**Q1. Categorize the 550 books to fiction and non-fiction.**

**Answer:**

i. select name, genre from amazon\_csv where genre = 'Fiction';

ii. select name, genre from amazon\_csv where genre = 'Non Fiction;

iii. SELECT

name AS Title,

author, genre,

row\_number() OVER (PARTITION BY genre) AS row

FROM Amazon\_csv;

iv. select count(case when genre='Fiction' then 1 end) Fiction,

count(case when genre='Non Fiction' then 1 end) NonFiction

from [dbo].[Amazon\_csv] group by genre;

**Observation :**

**Fiction count = 240**

**Non-Function = 310**

**Q2. Query the top 50 best sellers and show the following:**

**a. Title of Book**

**b. Author**

**c. Year**

**d. Review**

**Answer:**

select top(50) name, author, year,reviews from amazon\_csv order by reviews desc;

**Observation:**

**\* The range of top 50 reviews from 87841 to 26234**

**\* Year includes 2010 to 2019**

**\* Year excludes 2009**

**Q3. Query the books which have a rating greater than 4 released last 2019.**

**Answer:**

select name,user\_rating, year from amazon\_csv where user\_rating>4 and year=2019;

**Observation:**

**\* Total count rows have a rating >4 released in the year 2019=50**

**\* Average rating in the year = 4.74**

**Q4. Query the books which have greater than 10k reviews last 2018**

**Answer:**

select name,reviews, year from amazon\_csv where reviews>10000 and year=2018;

**Observation:**

**\* Total count rows have a reviews >10k released in the year 2018=26**

**\* Average rating in the year = 20616**

**Additional Queries:**

**1.Book count per year**

**Answer:**

select year, count(name) from amazon\_csv group by year;

**Observation:**

**\* Book count per year = 50 ( 2009 to 2019)**

**\* Total Book count = 550**

**\* Total year = 10**

**2. Top 5 reviews per year**

**Answer:**

select top(5) year, reviews from amazon\_csv order by reviews desc ;

**Observation:**

**\* In year 2019 = 87841**

**\* In year 2015 & 2016 = 79446**

**\* In year 2018 & 2019 = 61133**

**3. Top 5 reviews per user\_rating**

**Answer:**

select top 5 name, author, reviews, user\_rating from amazon\_csv order by user\_rating desc ;

**Observation:**

**\* Top first 2 reviews with 4.9 user rating = 14344**

**\* Top 3 reviews with 4.9 user rating = 7235**

**\* Top 3 reviews with 4.9 user rating = 5062**

**\* Top 4 reviews with 4.9 user rating = 4786**

**4. Highest price book in fiction**

**Answer:**

select max(price) from amazon\_csv where genre='Fiction';

**Observation:**

**\* Maximum price = 105**

**5. Lowest price, minimum user rating per year.**

**Answer:**

select year, min(price) as min\_price, min(user\_rating) as min\_user\_rating from amazon\_csv

group by year;

**Observation:**

**\* Minimum price from 2009 to 2019**

**Years 2010,2011,2013,2014,2015,2016,2017= 0**

**Year 2009 = 1**

**Year 2012, 2018,2019 = 10**

**\* Minimum User Rating from 2009 to 2019**

**The year 2012 = 3.3**

**The year 2015 = 3.6**

**The year 2013 = 3.8**

**The year 2014 = 3.9**

**Years 2009,2010,2011,2016,2017 = 4**

**Year 2018 = 4.2**

**Year 2019 = 4.3**