**Bookshop Analysis**

**Scenario:**

A new bookshop is opening in town and the owner would like an analysis of the most popular genres for fiction and non-fiction books. They would also like top rated books in those genres to display in the store to jump start sales.

**Data:**

The dataset that was used was the “Amazon 50 Top Bestselling Books from 2009-2019”. This data set shows the Name, Author, User Rating, Reviews, Price, Year, and Genre (fiction or non-fiction).

**Analysis:**

Below is the SQL used in BigQuery to clean and filter the data for fiction books. I wanted to add in specific genres other than fiction and nonfiction, so I filtered and saved a table to get rid of duplicates, only pull the column for Name and only the fiction genre. I then added the different genre types by conducting Google searches on the books and adding to the table. Once the table was complete, I uploaded into BigQuery and joined my new table with the original table to add the column “Genre\_Type”. I then used the COUNT and SUM functions to complete the analysis.

**Link to Final Table:**

[Final\_Books\_Fiction](https://docs.google.com/spreadsheets/d/1ReG6W8OmDaH_K9mneiRBG8WMJ1qtFO5S9IxvVH_cNvU/edit?usp=sharing)

**SQL Queries:**

# Running DISTINCT to get rid of duplicates

SELECT

DISTINCT Name

FROM `rare-nectar-344415.Best\_Books.books`

# Filtering Table to have only 3 columns needed for analysis and only fiction genre

SELECT

DISTINCT Name, User\_Rating, Reviews

FROM `rare-nectar-344415.Best\_Books.books`

WHERE

Genre = 'Fiction'

# Narrowing down data to top reviews and the most ratings

SELECT

DISTINCT Name, User\_Rating, Reviews

FROM

`rare-nectar-344415.Best\_Books.books`

WHERE

User\_Rating >= 3 AND Reviews >= 10000 AND Genre = ‘Fiction’

# Too many results, narrowing down more

SELECT

DISTINCT Name, User\_Rating, Reviews

FROM

`rare-nectar-344415.Best\_Books.books`

WHERE

User\_Rating >= 4 AND Reviews >= 10000

#Counting total genre to find most popular

SELECT

COUNT(Genre\_Type)

FROM `rare-nectar-344415.Best\_Books.Book\_Genres`

WHERE Genre\_Type = 'Dystopian'

#Dystopian = 11

SELECT

COUNT(Genre\_Type)

FROM `rare-nectar-344415.Best\_Books.Book\_Genres`

WHERE Genre\_Type = 'Romance'

#Romance = 6

SELECT

COUNT(Genre\_Type)

FROM `rare-nectar-344415.Best\_Books.Book\_Genres`

WHERE Genre\_Type = 'Historical'

# Historical = 11

SELECT

COUNT(Genre\_Type)

FROM `rare-nectar-344415.Best\_Books.Book\_Genres`

WHERE Genre\_Type = 'Fantasy'

# Fantasy = 9

SELECT

COUNT(Genre\_Type)

FROM `rare-nectar-344415.Best\_Books.Book\_Genres`

WHERE Genre\_Type = 'Childrens'

#Childrens = 10

SELECT

COUNT(Genre\_Type)

FROM `rare-nectar-344415.Best\_Books.Book\_Genres`

WHERE Genre\_Type = 'Mystery'

#Mystery = 6

SELECT

COUNT(Genre\_Type)

FROM `rare-nectar-344415.Best\_Books.Book\_Genres`

WHERE Genre\_Type = 'Thriller'

#Thriller = 7

# Top 3 Genres are Historical, Dystopian, and Childrens. These will be the genres for the displays.

# Joining tables to ratings, reviews, and genre

SELECT DISTINCT `rare-nectar-344415.Best\_Books.Reviews\_Ratings`.name, `rare-nectar-344415.Best\_Books.Reviews\_Ratings`.User\_Rating,`rare-nectar-344415.Best\_Books.Reviews\_Ratings`.Reviews, `rare-nectar-344415.Best\_Books.Book\_Genres`.Genre\_Type

FROM `rare-nectar-344415.Best\_Books.Reviews\_Ratings`

LEFT JOIN `rare-nectar-344415.Best\_Books.Book\_Genres`

ON `rare-nectar-344415.Best\_Books.Reviews\_Ratings`.name = `rare-nectar-344415.Best\_Books.Book\_Genres`.Name

#Query to group by and order by for final analysis

SELECT name, User\_Rating, Reviews, Genre\_Type

FROM `rare-nectar-344415.Best\_Books.Final\_Books`

WHERE Genre\_Type = 'Historical'

ORDER BY User\_Rating DESC

#Top 5 Historical Books to display are "The Help", "The Nightingale: A Novel", "To Kill a Mockingbird", "A Gentleman in Moscow: A Novel", and "All the Light We Cannot See"

SELECT name, User\_Rating, Reviews, Genre\_Type

FROM `rare-nectar-344415.Best\_Books.Final\_Books`

WHERE Genre\_Type = 'Dystopian'

ORDER BY User\_Rating DESC

# Top 5 Books for Dystopian "The Hunger Game" series, "1984", "Divergent" series, "Fahrenheit 451", and "Ready Player One: A Novel"

SELECT name, User\_Rating, Reviews, Genre\_Type

FROM `rare-nectar-344415.Best\_Books.Final\_Books`

WHERE Genre\_Type = 'Childrens'

ORDER BY User\_Rating DESC

#Top 5 Books for Childens "Brown Bear, Brown Bear, What Do You See?", "Last Week Tonight", "Oh, The Places You'll Go!", "The Very Hungry Caterpillar","Dear Zoo: A Lift-the-Flap Book"

#Running SUM on top 3 genres to which at the most reviews

SELECT

SUM(Reviews)

FROM `rare-nectar-344415.Best\_Books.Final\_Books`

WHERE Genre\_Type = 'Dystopian'

#Dystopian has 259,453 Reviews

SELECT

SUM(Reviews)

FROM `rare-nectar-344415.Best\_Books.Final\_Books`

WHERE Genre\_Type = 'Historical'

# Historical has 225,053 Reviews

SELECT

SUM(Reviews)

FROM `rare-nectar-344415.Best\_Books.Final\_Books`

WHERE Genre\_Type = 'Childrens'

# Childrens has 179,629 Reviews

#Dystopian was the most reviewed genre out of the top 3.

**Conclusion**

After completing the analysis of fiction books, the top 3 genres were Dystopian, Historical and Children’s. Dystopian was the most popular genre.

Below are the top 5 books to display out of each genre.

**Dystopian:**

* "The Hunger Game" series
* "1984"
* "Divergent" series
* "Fahrenheit 451"
* "Ready Player One: A Novel"

**Historical:**

* "The Help"
* "The Nightingale: A Novel"
* "To Kill a Mockingbird"
* "A Gentleman in Moscow: A Novel"
* "All the Light We Cannot See"

**Children’s**

* "Brown Bear, Brown Bear, What Do You See?"
* "Last Week Tonight"
* "Oh, The Places You'll Go!"
* "The Very Hungry Caterpillar”
* "Dear Zoo: A Lift-the-Flap Book"