Human Resources dataset analysis

*NAME: Sayantan Malladeb*

*CONTACT:* [*sm103@illinois.edu*](mailto:sm103@illinois.edu)

**Data Cleaning:**

* **Initial data attribute types**

A screenshot of a computer

Description automatically generated

* **Setting data types and their lengths   
  A screen shot of a computer

  Description automatically generated**

RESULT:

A screenshot of a computer

Description automatically generated

* **Adjusting Date values within dataset**
* A screenshot of a computer

  Description automatically generatedInitial birthdate values :

* Changing birthdate to one single format (YEAR/MONTH/DAY) : A screenshot of a computer

  Description automatically generated
* A screenshot of a computer

  Description automatically generatedConverting birthdate type from text to date:
* A screenshot of a computer

  Description automatically generatedInitial hire\_date values :
* Changing hire\_date to one single format (YEAR/MONTH/DAY) :

A screenshot of a computer

Description automatically generated

* A screenshot of a computer

  Description automatically generatedConverting hire\_date type from text to date:
* A screenshot of a phone

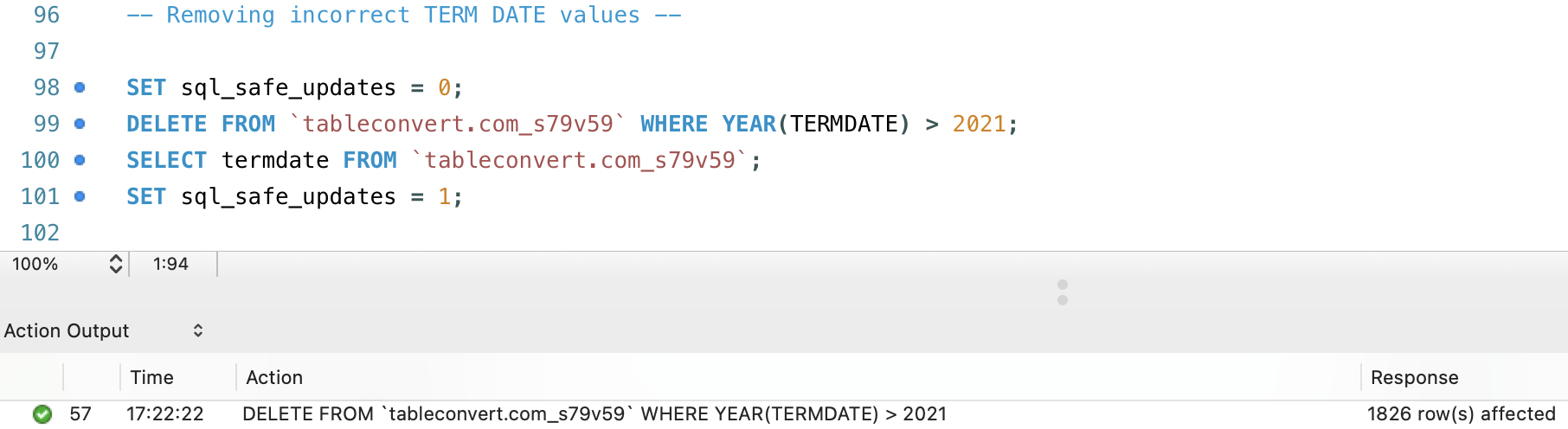
  Description automatically generatedInitial termdate (termination date) values :
* Removing unnecessary timestamp from termdate values: A screenshot of a computer

  Description automatically generated

A screenshot of a computer

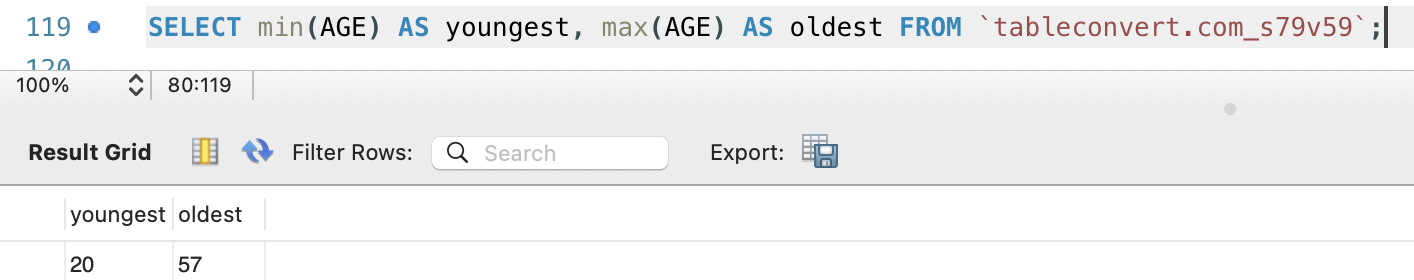
Description automatically generated

* Converting termdate type from text to date:
* Removing incorrect termdates (termdates greater than the year 2021):



* A screenshot of a computer

  Description automatically generatedAdding Age attribute:
* Removing incorrect negative age values and values <18:



**Questions for Analysis:**

* What are the top 5 most common job titles in the dataset?

A screenshot of a computer

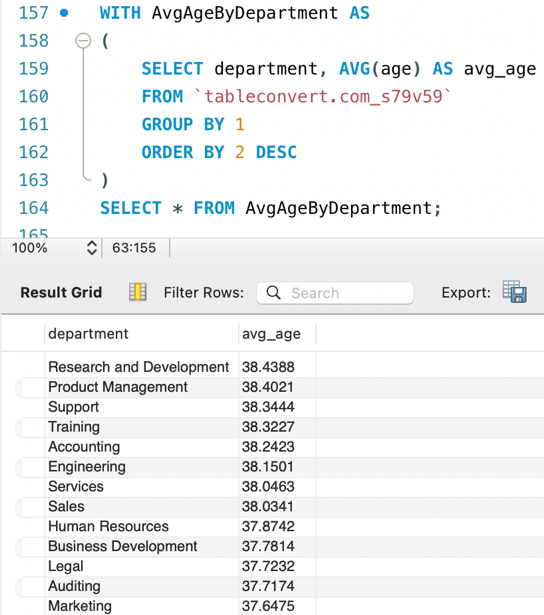
Description automatically generated

* Provide a breakdown of the distribution of races among employees for diversity and inclusion analysis?

A screenshot of a computer

Description automatically generated

* What is the average age of employees in each department?



* How many employees were hired in each year, and what's the trend?

A screenshot of a computer

Description automatically generated

* A screenshot of a computer

  Description automatically generatedWhat is the retention rate of employees based on gender?
* How many employees have been hired for each location over the years?

A screenshot of a computer

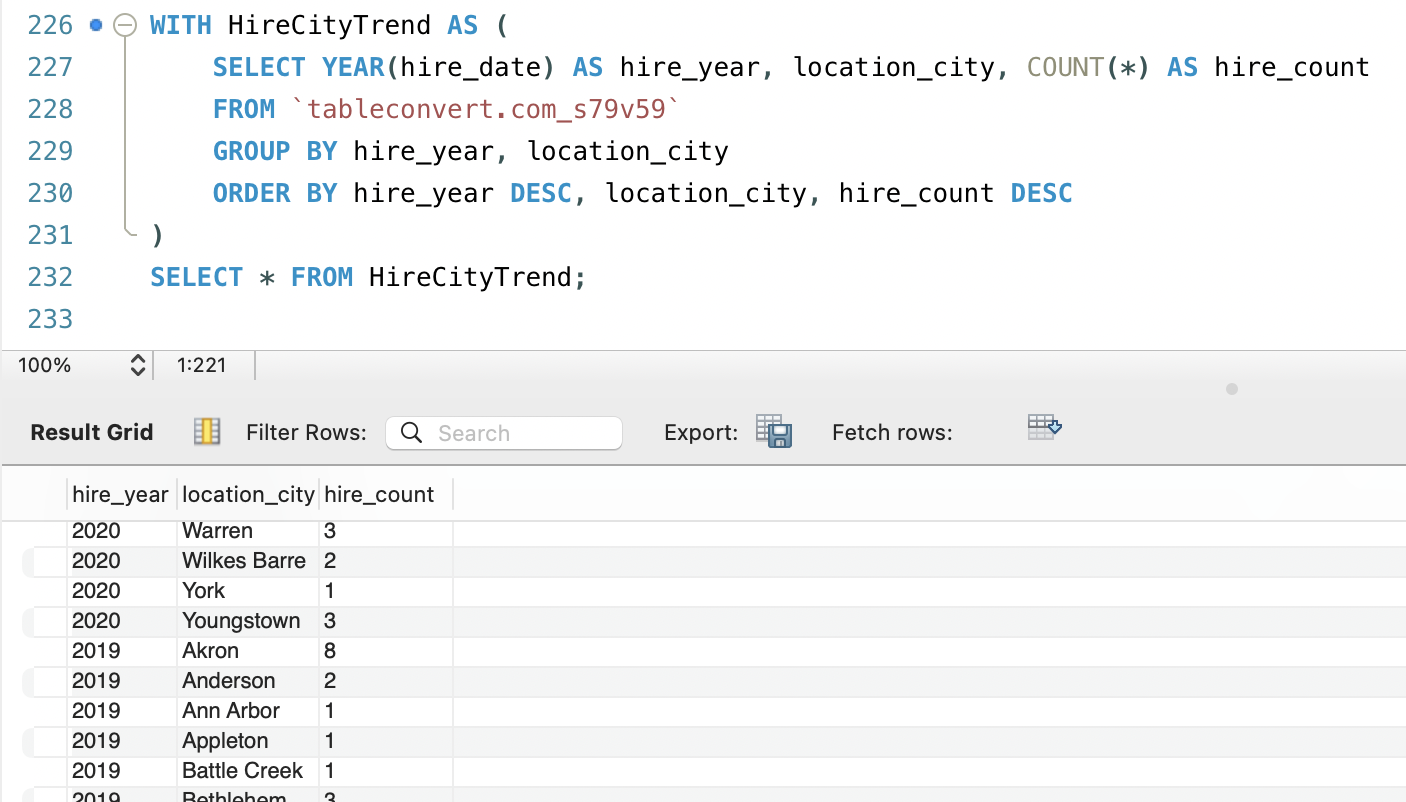
Description automatically generated

* What is the distribution of employees' ages by gender?

A screenshot of a computer

Description automatically generated

* How many employees have been hired in each city, and what's the trend?



* Identify the top 5 most common combination of department and job title?

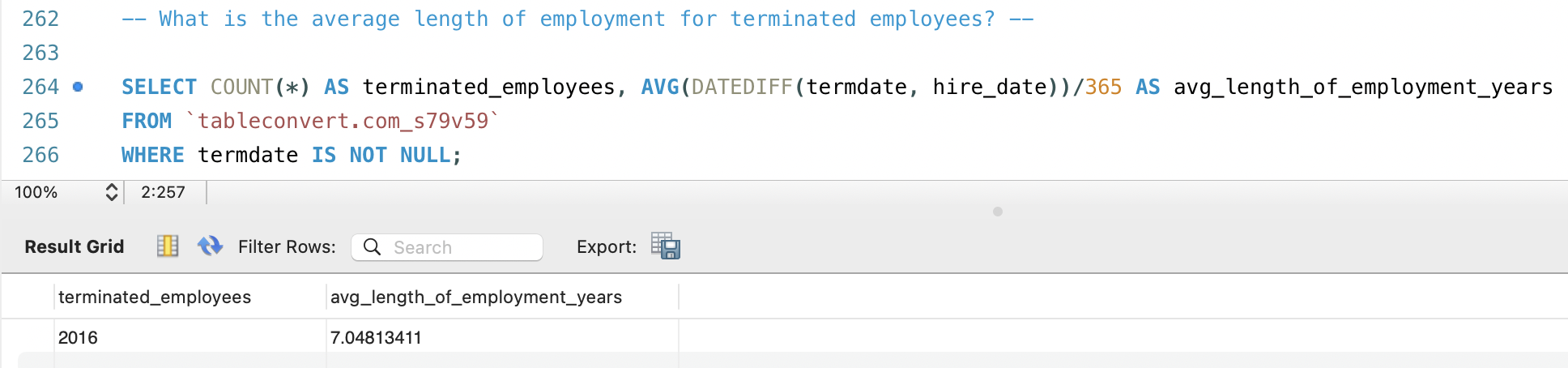
A screenshot of a computer

Description automatically generated

* How many employees have left each year?

A screenshot of a computer

Description automatically generated

* What is the average length of employment for terminated employees?
* Which departments have the top 5 highest turnover rate?

A screenshot of a computer

Description automatically generated

* What is the tenure distribution for each department?

A screenshot of a computer

Description automatically generated