S&P 500 Pay Ratio Report

Celine Dai | Chase Freund | Clint Krapfl | Yanni Mihalopoulos | Zach McDaniel

Introduction

CEO's in the Information Technology Industry are earning 8-figure salaries every year. (https://www.thestreet.com/investing/the-10-most-overpaid-ceos-running-multi-billion-dollar-companies-13045780) Some may argue that their payment is far higher than it should be. How does this lump salary compare to the work that is done by an average worker within the company? We want to compare the annual salary of CEOs (with stock options) to the average worker's annual salary. Potential users of the database can range from curious people in general to newspaper reporters and other fact-based professions.

Data

In this project, we are using AFL-CIO (https://aflcio.org/paywatch/highest-paid-ceos) to pull data about current Information Technology companies and their CEO's salary. Within the information technology field, we are looking at 67 rows of data. The original data set included all industry CEO salaries, but we have only pulled data referencing the information technology field. With that being said, we will be using a sample of approximately 100 companies and their CEO information.

Data Dictionary

Attribute	Type	Description
Ticker	Text	Unique ID for each company
Company Name	Text	Specific name for company in information technology industry
Company Type	Text	Type of corporation
State ID	Text	Unique ID that identifies each state
State Name	Text	Full name of each state in the United States
City Name	Text	Full name of each city per state
CEO Name	Text	First, middle, last, and suffix of each CEO
CEO Pay	Numeric	Specific pay per CEO annually
Pay Year	Text	Year that payment information was recorded
CEO Pay	Numeric	Sum of salary and stock options
Median Pay	Numeric	Average annual worker pay across entire company
Pay Ratio	Text	The ratio of CEO pay to average worker pay annually
Revenue	Numeric	Total annual revenue accrued per company (performance metric)

ERD Description

The primary entity in the database is COMPANY, which is identified by TickerID. The raw data contains many composite attributes that can be modeled as weak entities. In our case, a single company has a single CEO and financial report. In addition, all business rules from COMPANY are M:M. For example, COMPANIES have one and only one HEADQUARTERS. Figure 1 displays the ERD model for this data.

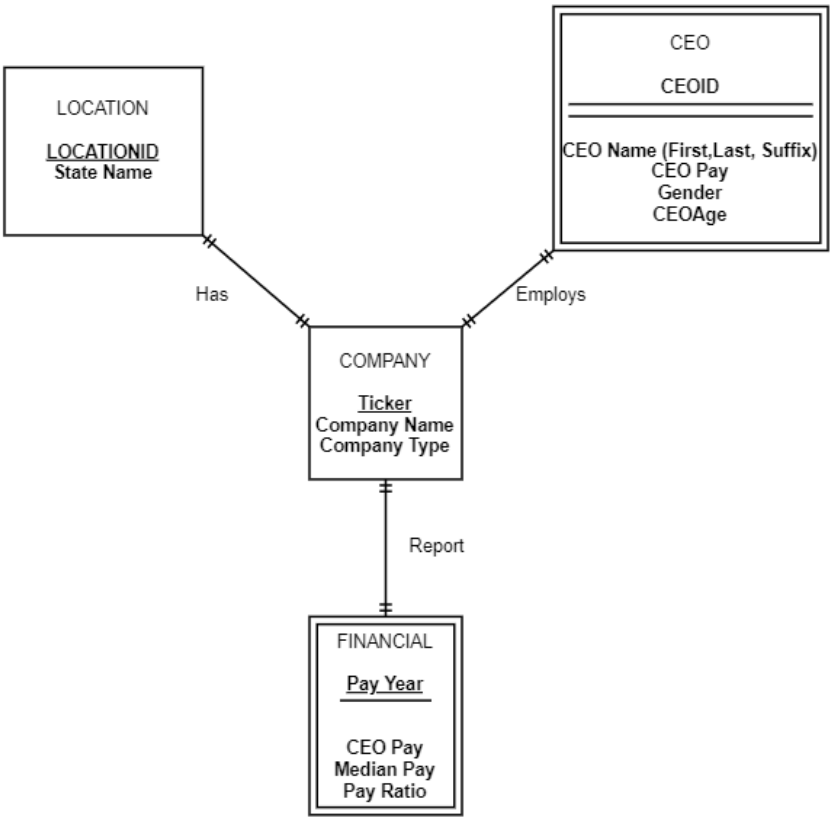

Figure 1: Entity Relationship Diagram

Figure two displays our graphical relational schema. In this schema, we created four total tables. Of which, COMPANY acts as our parent table and each child table holds the TickerID or LocationID as the foreign key.

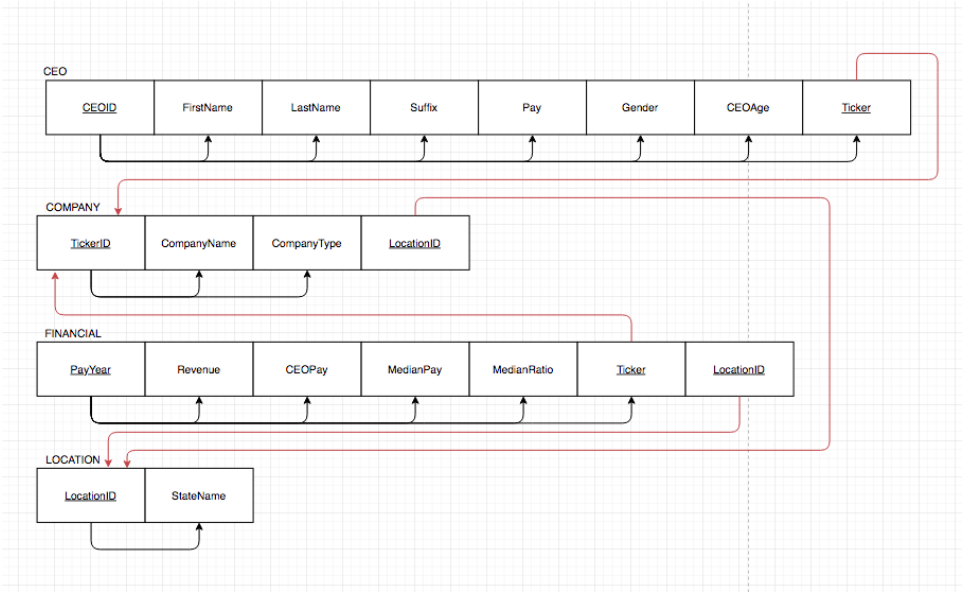

Figure 2: Graphical Relational Schema

Database Implementation

In Oracle Apex, we wrote CREATE TABLE for the desired tables in our ERD.

LOCATION

The location table was created next along with the following:

```
CREATE TABLE "LOCATION" (
LocationID VARCHAR2(5) NOT NULL,
StateName VARCHAR2(10) NOT NULL,
CONSTRAINT Location_PK PRIMARY KEY (LocationID) );
```

COMPANY

Company is our parent table for all tables other than location

```
CREATE TABLE "COMPANY" (
```

TickerID VARCHAR2(5) NOT NULL,

CompanyName VARCHAR2(60) NOT NULL,

CompanyType VARCHAR2(8),

LocationID VARCHAR2(2),

CONSTRAINT Ticker_PK PRIMARY KEY (TickerID),

CONSTRAINT Ticker_FK FOREIGN KEY (LocationID) REFERENCES LOCATION (LocationID));

CEO

CREATE TABLE "CEO" (

CEOID VARCHAR2(15) NOT NULL,

CEOFirst VARCHAR2(15) NOT NULL,

CEOLast VARCHAR2(15) NOT NULL,

CEOSuffix VARCHAR2(15),

CEOPay NUMBER(8),

CEOAge NUMBER(2) NOT NULL,

TICKER VARCHAR2(4),

GENDER VARCHAR2(6),

CONSTRAINT CEO_PK PRIMARY KEY (CEOID),

CONSTRAINT CEO_FK FOREIGN KEY (TICKER) REFERENCES COMPANY (TICKERID)) ;

FINANCIAL

CREATE TABLE "FINANCIAL" (

PayYearID VARCHAR(2) NOT NULL,

PayYear CHAR(4) NOT NULL,

CEOPay NUMBER(9) NOT NULL,

MedianPay NUMBER(6) NOT NULL,

PayRatio VARCHAR2(4) NOT NULL,

TICKER VARCHAR2(4) NOT NULL,

LocationID VARCHAR2(2) NOT NULL,

CONSTRAINT PayYearID_PK PRIMARY KEY (PayYearID),

CONSTRAINT PayYearID_FK FOREIGN KEY (TICKER) REFERENCES COMPANY (TICKERID),

CONSTRAINT PayYearID_FK2 FOREIGN KEY (LocationID) REFERENCES LOCATION (LocationID)) ;

Analysis

Our analysis is intended to highlight interesting aspects of the S&P 500 information technology CEO's financial information. The goal is to better understand pay disparity in different aspects of these companies such as gender differences, median worker pay differences, and general differences among the CEOs.

Question 1:

Our goal is to identify which states and which companies have the highest pay ratios. The following report lists the top 5 companies with the highest pay ratios, and which states they reside in. To address this question, we composed a simple query that averages the pay ratio of each company and ordered them by descending order from highest to lowest average pay ratios.

```
SELECT LOCATION.STATENAME, AVG(FINANCIAL.PAYRATIO) AS AVERAGERATIO
FROM LOCATION, FINANCIAL
WHERE LOCATION.LOCATIONID = FINANCIAL.LOCATIONID
GROUP BY LOCATION.STATENAME
ORDER BY AVERAGERATIO DESC;
```

The results of this query are shown below (Figure 3). From a large selection pool, we figured that CT had the highest pay ratios while WA had the lowest pay ratios from our data set. These results for information technology industry are consistent with current trends, as states with many companies in this industry have lower pay ratios due to proximity to other competing businesses.

Statename	Averageratio
CT	437
NJ	412
AZ	313
CO	306
FL	253
ID	252
GA	225
IL	211
TX	210
NY	204
CA	190
VA	189
MA	188
NC	179
OH	159

Figure 3: Highest Pay Ratios

Because of a multitude of factors playing into pay ratio, we wanted to measure the top 5 highest pay ratio companies in the information technology field. We decided to measure these companies to see how they compare to our previous query which returned average pay ratio by state.

```
SELECT TICKERID ||', '|| COMPANYNAME AS COMPANY, PAYRATIO
FROM COMPANY, FINANCIAL
WHERE COMPANY.TICKERID = FINANCIAL.TICKER
ORDER BY CAST(PAYRATIO AS INT) DESC
FETCH FIRST 5 ROWS ONLY;
```

Company	Payratio
WDC, Western Digital	1279
APH, Amphenol	766
PYPL, PayPal Holdings	543
SWKS, Skyworks Solutions	447
DXC, DXC Technology	415

Figure 4: Average Pay Ratios by State

Question 2:

Our goal is to identify the CEOs who make significant money with relatively lower pay ratio and younger age. The following report lists the name of CEO's who makes over 10 million dollars followed by pay ratio, and CEO age. To find the data we created a simple query that selects the CEO pay and pay ratio comes with specific CEO names.

.

```
SELECT CEO.CEOFIRST ||' '|| CEO.CEOLAST AS NAME, CEO.CEOPAY, FINANCIAL.PAYRATIO
FROM CEO, FINANCIAL
WHERE CEO.CEOPAY = FINANCIAL.CEOPAY
AND CEO.CEOPAY >= 10000000
AND FINANCIAL.PAYRATIO <= 150 ;
```

The results of this query are shown below (Figure 5). Form this query, we figured out all the CEOs who make over 10 million dollars and have a pay ratio less than 150.

NAME	CEOPAY	PAYRATIO
Brad Smith	21071738	143
Alfred Kelly	19493946	147
David Henshall	19258102	113
Gary Dickerson	14064540	123
Jen-Hsun Huang	13642838	88
George Kurian	13164969	67
Eugene Hall	11502517	107
F. Leighton	11347676	103
Ronald Nersesian	10240253	89
Ajei Gopal	10164184	71

Figure 5: CEO Pay and Pay Ratio

Since we also interested in the relationship between CEP pay and their ages, we figured out all the CEOs who make more than 10 million dollars at an age under 55.

```
SELECT CEO.CEOFIRST ||' '|| CEO.CEOLAST AS NAME, CEO.CEOPAY, CEOAGE
FROM CEO, FINANCIAL
WHERE CEO.CEOPAY = FINANCIAL.CEOPAY
AND CEO.CEOPAY >= 10000000
AND CEOAGE<=55
ORDER BY CEO.CEOPAY DESC;
```

NAME	CEOPAY	CEOAGE
Marc Benioff	28391846	55
Satya Nadella	25843263	52
Charles Robbins	21284339	53
Steven Mollenkopf	19975472	51
David Henshall	19258102	51
Dion Weisler	19215534	52
Carlos Rodriguez	19000187	55
Gary Norcross	18451529	53
Jeffrey Sloan	16818560	52
Francisco D'Souza	14094531	51

Figure 6: CEO Pay and Age

Question 3:

Our goal is to identify the discrepancies between Male and Female CEO pay. We want to measure this to provide a simplistic visualization pay differentials and argue against or for such wage gaps. To address this question, we created a complex query (including a subquery) to measure the total female CEO pay compared to the overall CEO pay of both genders. Below is a classic report in which the total Male vs. Female CEO pay is listed in percentages to the whole.

```
SELECT GENDER, COUNT(GENDER) AS TOTALCEO, ROUND(SUM(CEOPAY)
/(SELECT SUM(CEOPAY)
FROM CEO) * 100)  || '%'
AS PERCENTAGE_OF_TOTAL
FROM CEO
GROUP BY GENDER
ORDER BY PERCENTAGE_OF_TOTAL DESC;
```

The results of this complex query are shown below (Figure 7). This query gave expected results as the total amount of Females represented in our Data Set is 4. This leads to further questioning on why Females are not regularly CEO's of the Technology Information industry. Our findings show a 90% disparity between the gender's salary.

Gender	Totalceo	Percentage Of Total
Male	57	95%
Female	4	5%

Figure 7: Gender CEO Pay

With this information, and query results, we can now look deeper into median pay ratios of each company and their locations which may give insight to more findings.

Question 4:

To further understand the relation between locations and the pay ratios we want to identify which states pay their employees the best. To measure this, we identified a median pay of $100,000 as our mark. To address this question, we wrote a query that returned a list of all states that have companies whose median pay was equal to or greater than $100,000; next, we found the quantity of companies having that median pay per each state to better illustrate which states have higher pay.

```
SELECT LOCATION.STATENAME, COUNT(FINANCIAL.LOCATIONID) AS QUANTITY
FROM LOCATION JOIN FINANCIAL
ON LOCATION.LOCATIONID = FINANCIAL.LOCATIONID
WHERE MEDIANPAY > 100000
GROUP BY LOCATION.STATENAME
ORDER BY QUANTITY DESC;
```

Statename	Quantity
CA	17
MA	2
FL	1
VA	1
PA	1
NY	1
WA	1
CT	1

Figure 8: State's Quantity of $100,000 Median Pay or Higher

We represented this data in a bar chart visualization. This allows users to see how each company whose median pay is over $100,000 and to which state these companies belong to.

Web Design

Link to Our Application:

https://apex.oracle.com/pls/apex/dbm_finalproject/r/final-project4/home?session=12519728189653

Home Page:

On our homepage we added a brief description of our goal in the project and a custom navigation menu to organize our analysis questions 1 through 4. Figure 9 shows an image of our home page.

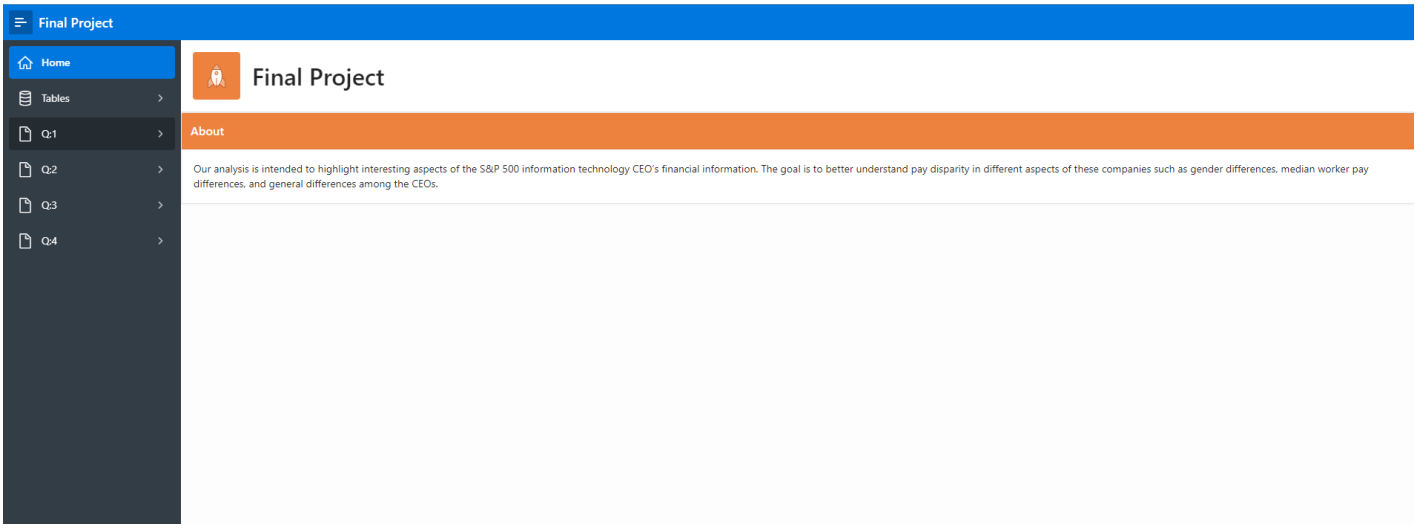

Figure 9: Home Page

Tables:

We created an interactive report for each database table, which allows users to search, filter, and group the data (Figures 10-13)

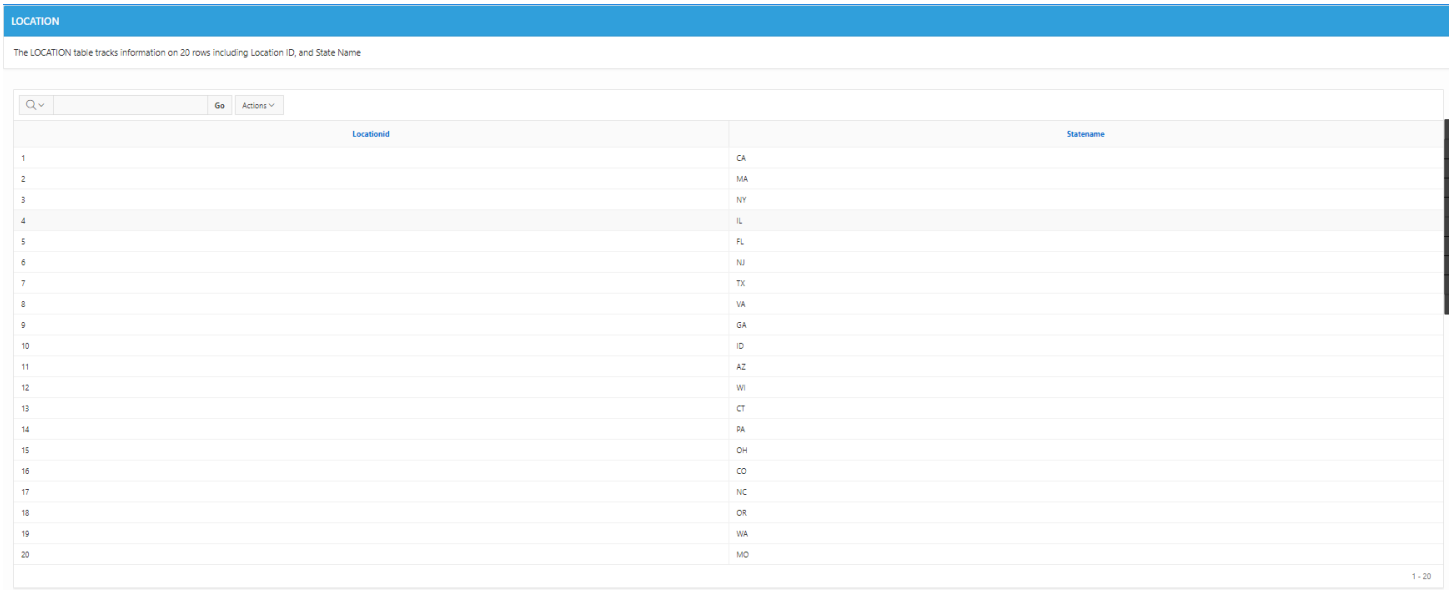

Figure 10: LOCATION Table

COMPANY

The COMPANY table tracks 61 rows including Ticker, Company Name, Company Type, and Location ID

Go Actions

Tickerid	Companyname	Companytype	Locationid
PYPL	PayPal Holdings	Inc.	1
ADBE	Adobe	Inc.	1
CRM	salesforce.com	Inc.	1
MSFT	Microsoft	Corp.	2
CSCO	Cisco Systems	Inc.	1
INTU	Intuit	Inc.	1
KLAC	KLA	Corp.	1
MA	Mastercard	Inc.	3
MSI	Motorola Solutions	Inc.	4
QCOM	QUALCOMM	Inc.	1
V	Visa	Inc.	1
CTXS	Citrix Systems	Inc.	5
HPQ	HP	Inc.	1
ADP	Automatic Data Processing	Inc.	6
FIS	Fidelity National Information Services	Inc.	5
TXN	Texas Instruments	Inc.	7
IBM	International Business Machines	Corp.	3
SYMC	Symantec	Corp.	1
DXC	DXC Technology	Co.	8
GPN	Global Payments	Inc.	9
AAPL	Apple	Inc.	1
GLW	Corning	Inc.	3
MU	Micron Technology	Inc.	10
CTSH	Cognizant Technology Solutions	Corp.	6
AMAT	Applied Materials	Inc.	1
NVDA	NVIDIA	Corp.	1

Figure 11: COMPANY Table

FINANCIAL

The FINANCIAL table contains 61 rows with information including Ticker ID, Pay Year ID, Pay Year, CEO Pay, Median Pay, Pay Ratio, and Location ID

Go Actions

Payyearid	Payyear	Ceopay	Medianpay	Payratio	Ticker	Locationid
1	2018	37764588	69600	543	PYPL	1
2	2018	28397528	142192	200	ADBE	1
3	2019	28391846	151955	187	CRM	1
4	2018	25843263	167689	154	MSFT	2
5	2018	21284339	132764	160	CSCO	1
6	2018	21071738	147184	143	INTU	1
7	2019	20512367	85968	239	KLAC	1
8	2018	20379353	121897	167	MA	3
9	2018	20348558	96553	211	MSI	4
10	2018	19975472	85592	233	QCOM	1
11	2018	19493946	132483	147	V	1
12	2018	19258102	170433	113	CTXS	5
13	2018	19215534	79719	241	HPQ	1
14	2019	19000187	63225	412	ADP	6
15	2018	18451529	46929	393	FIS	5
16	2018	17576224	83905	210	TXN	7
17	2018	17552218	55088	319	IBM	3
18	2018	17347581	102869	169	SYMC	1
19	2019	17256539	41602	415	DXC	8
20	2018	16818560	60931	276	GPN	9
21	2018	15682219	55426	283	AAPL	1
22	2018	14921071	52093	286	GLW	3
23	2018	14241583	56540	252	MU	10
24	2018	14094531	34183	412	CTSH	6
25	2018	14064540	113999	123	AMAT	1
26	2019	13642838	155035	88	NVDA	1

Figure 12: FINANCIAL Table

CEO

The CEO table tracks information on 61 rows including Ceoid, Ceofirst, Ceolast, Ceosuffix, Ceopay, Ceoage, Ticker, Gender.

Go Actions

Ceoid	Ceofirst	Ceolast	Ceosuffix	Ceopay	Ceoage	Ticker	Gender
1	Daniel	Schulman		37764588	62	PYPL	Male
2	Shantanu	Narayen		28397528	56	ADBE	Male
3	Marc	Benioff		28391846	55	CRM	Male
4	Satya	Nadella		25843263	52	MSFT	Male
6	Charles	Robbins		21284339	53	CSCO	Male
7	Brad	Smith		21071738	61	INTU	Male
8	Richard	Wallace		20512367	71	KLAC	Male
9	Ajay	Banga		20379353	60	MA	Male
10	Gregory	Brown		20348558	60	MSI	Male
11	Steven	Mollenkopf		19975472	51	QCOM	Male
12	Alfred	Kelly	Jr.	19493946	61	V	Male
13	David	Henshall		19258102	51	CTXS	Male
14	Dion	Weisler		19215534	52	HPQ	Male
15	Carlos	Rodriguez		19000187	55	ADP	Male
16	Gary	Norcross		18451529	53	FIS	Male
17	Richard	Templeton		17576224	61	TXN	Male
18	Virginia	Rometty		17552218	62	IBM	Female
19	Gregory	Clark		17347581	62	SYMC	Male
20	J.	Lawrie		17256539	66	DXC	Male
21	Jeffrey	Sloan		16818560	52	GPN	Male
22	Timothy	Cook		15682219	59	AAPL	Male
23	Wendell	Weeks		14921071	61	GLW	Male
24	Sanjay	Mehrotra		14241583	61	MU	Male
25	Francisco	D'Souza		14094531	51	CTSH	Male
26	Gary	Dickerson		14064540	64	AMAT	Male
27	Jen-Hsun	Huang		13642838	57	NVDA	Male

Figure 13: CEO Table

Queries:

To answer question one, our query takes in data from the Location table and the Financial table to create a report for the average pay ratio. In the first column states with the highest ratio are put in order, in column two the corresponding pay ratio for that state is listed.

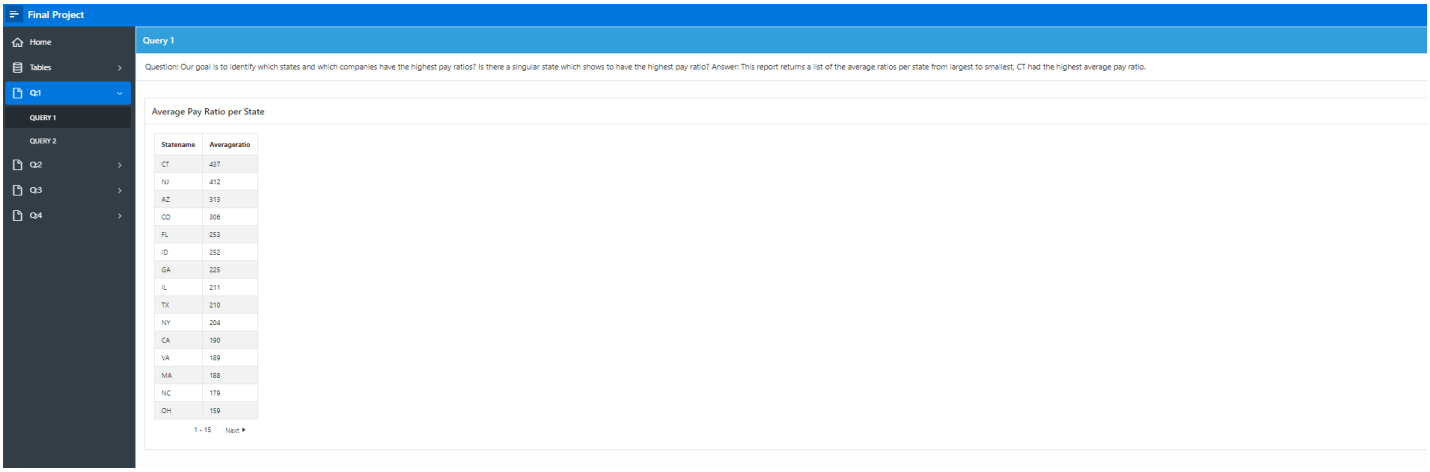

Figure 14: Query 1 Average Ratio Per State

As a follow up to query one, we created a second query that is less broad of a statistic and breaks down the pay ratio by the individual company instead of by state. Here you can see we decided to display only the top 5

companies in our data with the highest pay ratio (Ratio between CEO pay, and employees pay within that company)

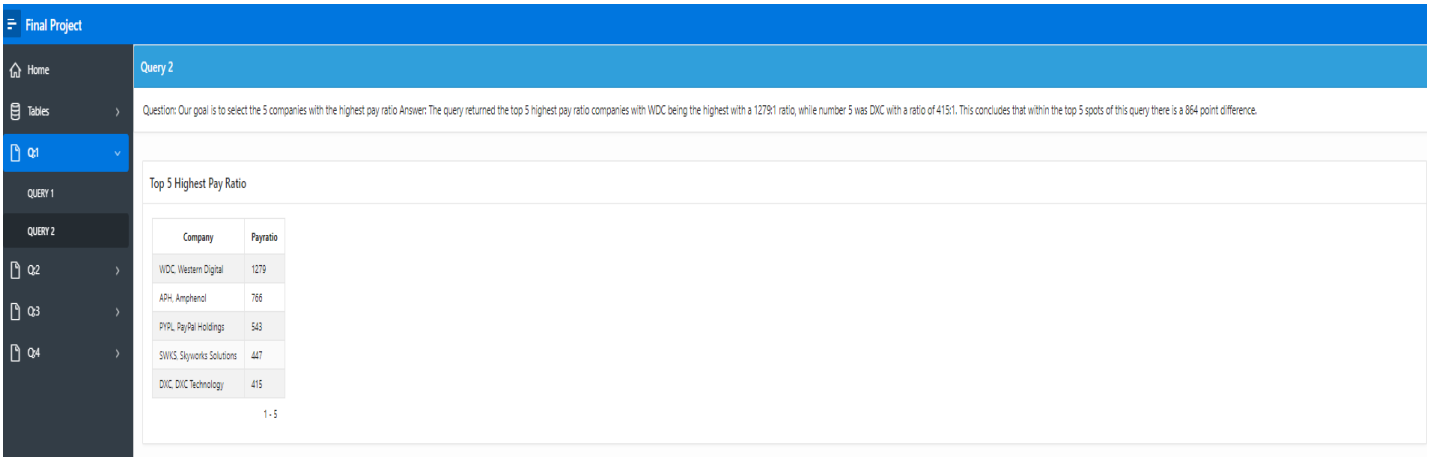

Figure 15: Top 5 Pay Ratios

To get a better idea of how much more CEO's are paid when compared to their employees, we created a report showing the CEO, company revenue, and pay ratio for that company. CEO's that make over $10,000,000 and have a pay ratio lower than 150:1 listed.

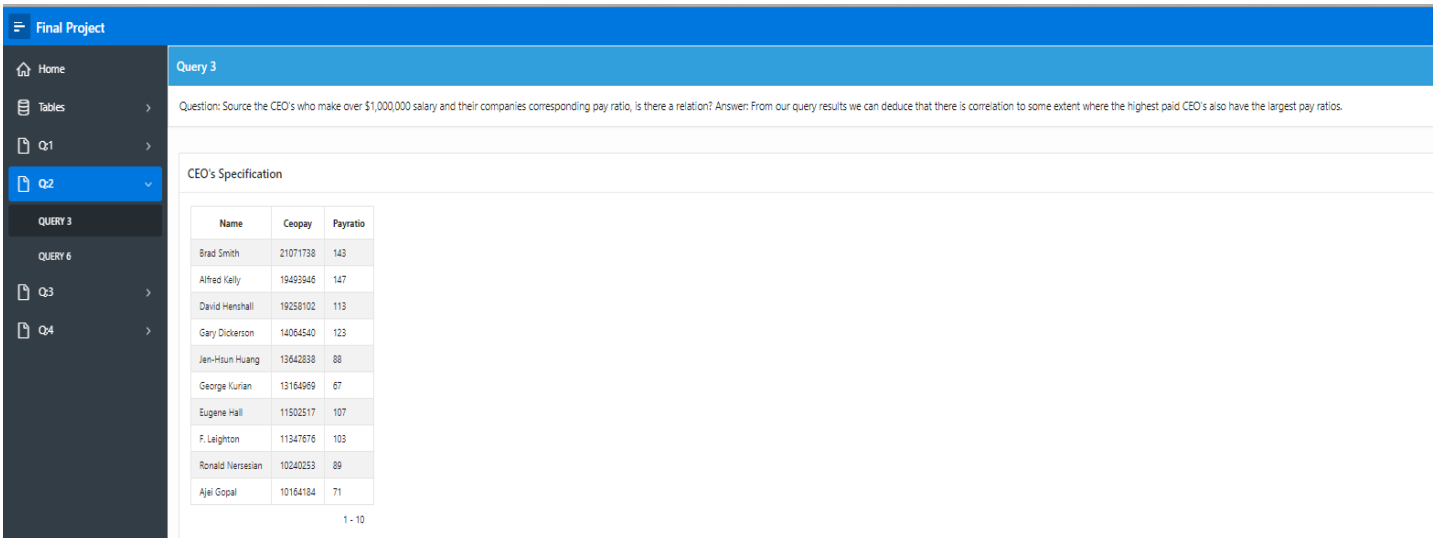

Figure 16: CEOs Over $10,000,000 with Pay Ratio Less Than 150

To answer question 3, we made a report that can show how age is correlated with pay. As you can see, age does not determine a vast difference in the compensation a CEO may receive.

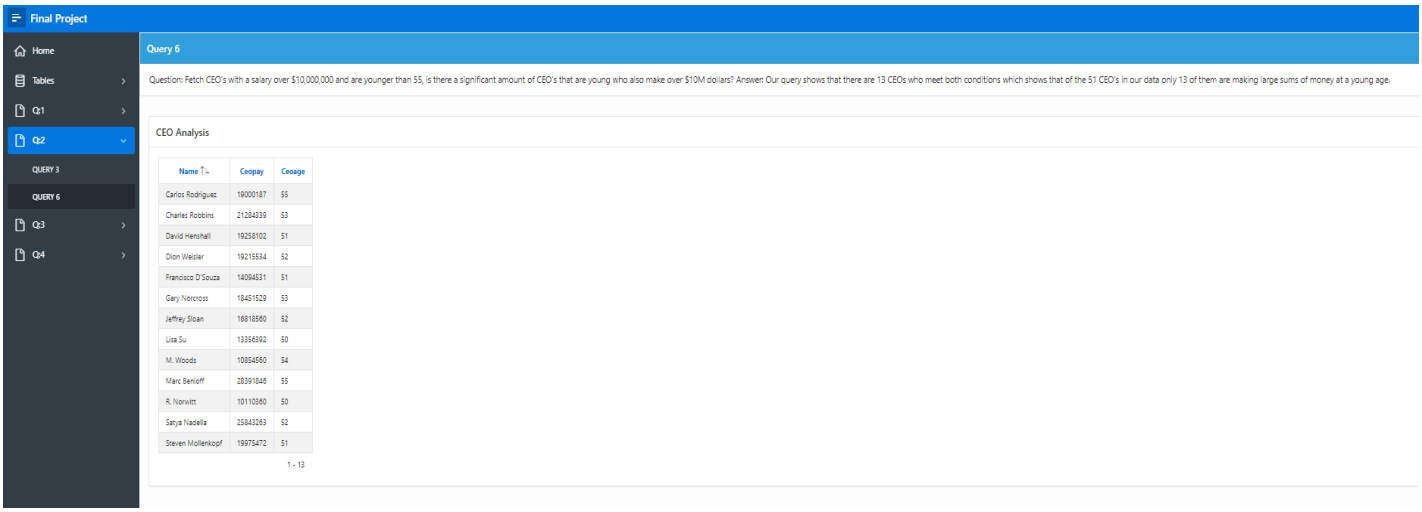

Figure 17: CEOs Over $10,000,000 Aged 55 or less

As a follow up to question 3, we asked what the pay distribution would be by gender. The report displays a comparison between our male data points and our female data points. To understand the comparison better, we pinned a pie chart to the right of the report for visualization purposes.

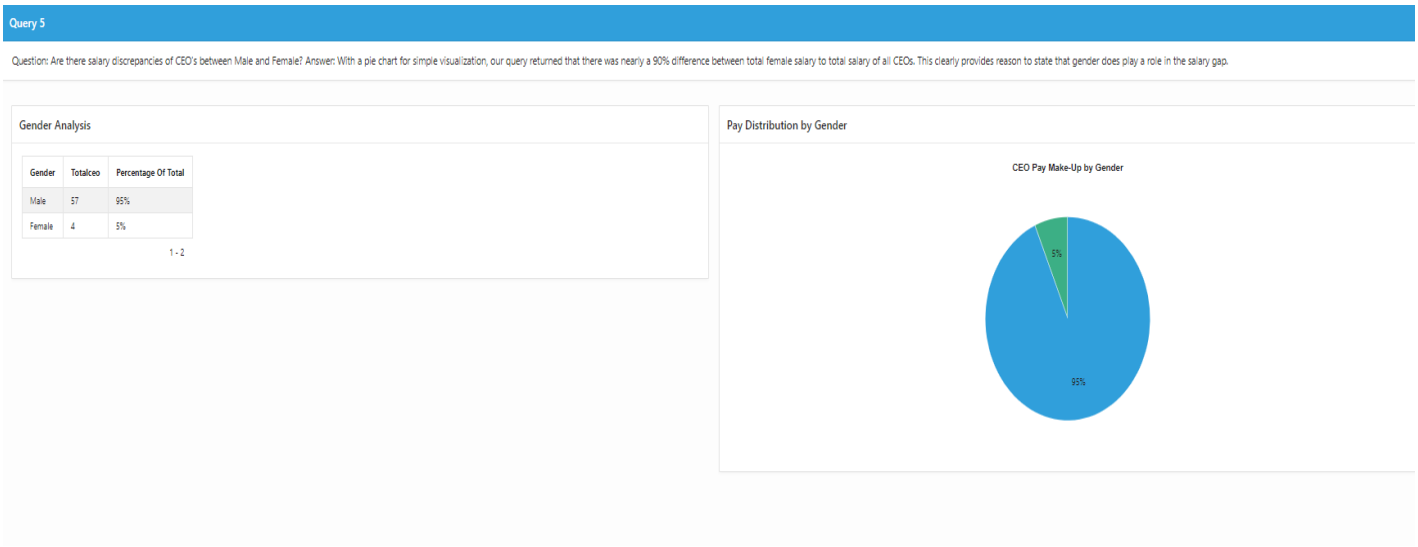

Figure 18: CEO Pay Distribution by Gender

Going off question 4, we created a query to find which companies in our data had a median pay over $100,000. We then took that data and grouped it by state, to see which states have higher median paying salaries. A bar chart was created to visually see which states had the most companies with median pay over $100,000.

Figure 19: Median Pay Over $100,000 by State