

MSBX 5405 Structured Data Modeling & Analysis Final Project Report

Submitted By: Alex Qaddourah, Alyson Chen, Junji Wiener, Wenbin Yang

Instructor: Keith Sparling

Project Scenario: Project Description

We started our analysis by using companies' stock price data sets. Our companies' industry include Software Infrastructure, Semiconductors, Internet Retail, Consumer Electronics, Consumer Electronics, Information Technology Services and Internet Content & Information. Moreover, those companies are large cooperation across the United States and also in China. As a future analyst and global citizen, knowing the stock prices' differences among those companies can help us improve the knowledge in the international stock market and also help us to learn more about the stock prices in the different industry. Additionally, we have two data sets which are the profile of those executives who work for the high-tech companies and another data set related to the detailed information of those companies such as HQ's address, phone number and total number of employees. We also camp up some interesting business questions for our project.

Project Scenario: Tables' Relationships

After creating the ER model, we found that all of the data set are related to 'all_profiles'. Before we started our project, we did not know there is a significant correlation among those data set. In sum, we think an ER model is a powerful tool and is a primarily tool to help communicate with data analysts on the project. Besides, it gives diagrammatical and graphical of various entities also its attributes and relationships between entities. As a result, the ER model helps us to understand the data structure and in minimizing redundancy and other problems.

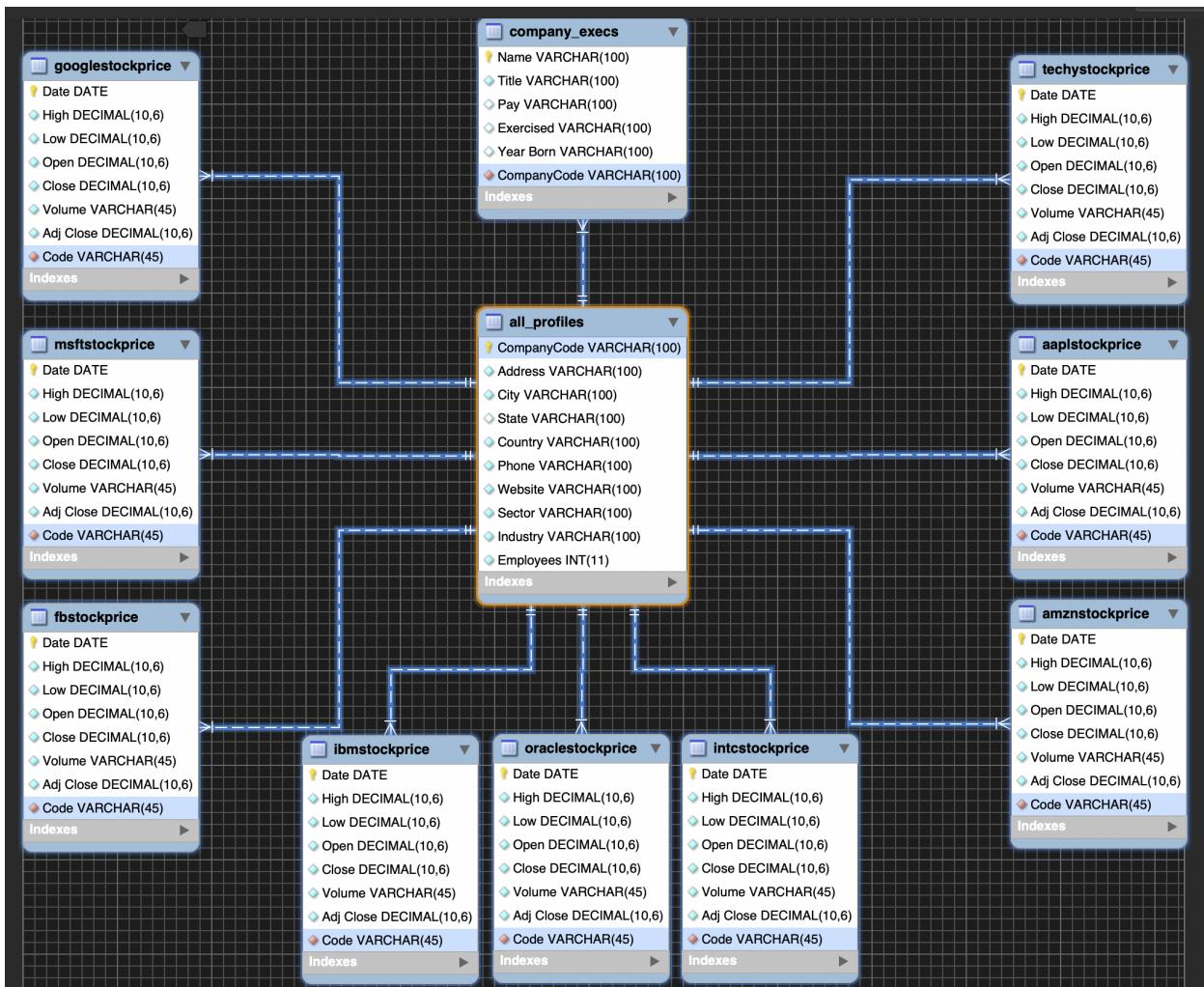
Project Scenario: Important Field and Columns

As we can see, there are nine stock price tables with the same columns, and we are more interested in the adj_close price for each company. There are two different ways to determining the value of a stock, namely the "closing price" and "adjusted closing price". Before we go deeper into closing price, we want to have better understanding about closing price and according to Bischoff, he stated "The closing price is simply the cash value of that specific piece of stock at day's end while the adjusted closing price reflects the closing price of the stock concerning other stock attributes. In general, the adjusted closing price is considered to be a more technically accurate reflection of the true value of the stock." (Bea Bischoff,2019)

As for 'all_profiles' and 'company_execs' we prefer to analyze the companies' location which not in United State and if it is in the States which state they are located in.

Besides, we are curious about the CEOs' salary and how many job opportunities are these companies provide for the world. Moreover, when it comes to data interception, we thought about the case that Diego told us" After the weekend people go into Wall Street on Monday feeling gloomy and as a consequence returns should be lower." Therefore, we want to testify the result by using MySQL. Since it's a huge project, we have to use create view to make the code easier to check.

ER Model



Select Statements

- 1. To see which companies' executives' are below 45 year-old.**

I want to know which companies' executives tend to be 'younger'. Originally, I set the age of executive was 35 but only found one result which was Mark Zuckerberg. Therefore, I set the age up to 45 years old.

- 2. To show the executives information at the company has employees more than 300000.**

To practice where function.

- 3. To show each company's close price on my 21st Birthday (2016-03-28).**

I am curious how much was each company's close price on my birthday back in 2016.

- 4. Executives who earns more than 10 million.**

I am curious about how many can those large cooperation executives make each year.

- 5. Display name,title,yearborn,company-code, pay for all executives who earn more than average pay and who work at Google.**

To practice my ability in building subqueries.

- 6. To compare the close price between Facebook and Tencent in 2015.**

To see what were the close price difference between Facebook and Tencent in 2015.

- 7. Display all the executives' names and title who work at Tencent**

It's the only Chinese cooperation among those companies, as a Chinese, it's quite interesting for me to know who and what do they do at Tencent.

- 8. Create a query to connect 'all_profiles' and 'company_execs'.**

To practice the basic query by using joins.

- 9. To show all the CEO's company-code, name and year-born from technology companies.**

To select all the title with 'CEO' and they are all from technology by using subqueries.

10. To see where those companies' location (by state) and which states' companies provided the most job opportunities.

I am curious which states have the most companies' HQ and also want to know what companies provided the most job opportunities.

11. Robert Shiller, the famous economist from Yale, is arguing with Eugene Fama, another famous economist, regarding the stock market. Eugene believes the market is pretty efficient, so that returns on any given day are unrelated to any other day. Robert thinks Mondays are different: after the weekend people go into Wall Street on Monday feeling gloomy, and as a consequence returns should be lower. According to this, Is the stock average close price lowest on Monday?

The conclusion is among those nine companies' average close price in Monday, We can say Investors are less active on Monday. This question is from one of our homework question in our Quantitative Methods class.

We used R to testified this question and I want to use MySQL to see if the result is the same.

12. Get the gap(max(high) - min(low)) of each company in each year.

I want to know the gap between all the companies in every year.

Select Statement Screenshot Area

Select Statement #1 #2

```
BusTartarus-FL19.ad.colorado.edu - _weya9345_2exampractice
1 use _weya9345_project;
2
3 #1 To see which companies CEO are below 45 year-old
4 select name, title, year(now())-YearBorn 'age', CompanyCode
5 from company_execs
6 where year(now())-YearBorn <= 45;
7
8 #2 To show the CEO information at the company has employees more than 300000.
9 select *
10 from company_execs
11 where CompanyCode in (
12 select companycode
13 from all_profiles
14 where Employees > 300000);
15 |
```

name	title	age	CompanyCode
Mr. Mark Elliot Zuckerberg	Founder, Chairman & CEO	35	FB
Mr. Michael T. Schroepfer	Chief Technology Officer	44	FB
Mr. Navin Shenoy	Exec. VP & GM of Data Platforms Group	45	INTEL
Mr. Sergey Brin	Co-Founder & Director	45	GOOG

Name	Title	Pay	Exercised	YearBorn	CompanyCode
Mr. Andrew R. Jassy	Chief Exec. Officer of Amazon Web Services Inc.	266232	1968	AMZN	
Mr. Brian T. Olsavsky	Sr. VP & CFO	163200	1964	AMZN	
Mr. Jeffrey A. Wilke	Chief Exec. Officer of Worldwide Consumer	255613	1967	AMZN	
Mr. Jeffrey M. Blackburn	Sr. VP of Bus. Devel.	178500	1970	AMZN	
Mr. Jeffrey P. Bezos	Founder, Chairman, Pres & CEO	1681840	1964	AMZN	
Dr. John E. Kelly III	Exec. VP	1730936	1954	IBM	
Mr. Fletcher Previn	Chief Information Officer			IBM	
Mr. James J. Kavanagh	Sr. VP & CFO	1609951	1967	IBM	
Mr. Martin J. Schroeter	Sr. VP of IBM Global Markets Global Financing, Marketing & Communications	2149949	1964	IBM	
Ms. Virginia M. Rometty	Chairman, Pres & CEO	6750826	1959	IBM	

Select Statement #3

```
BusTartarus-FL19.ad.colorado.edu - _weya9345_2exampractice
1 use _weya9345_project;
2
3 #3 To show each company's close price on my 21st Birthday (2016-03-28)
4 Select aa.Date, aa.close as 'apple', am.close as 'amazon', fb.close as 'Facebook', gg.close as 'Google',
5 ibm.close as 'IBM', ints.close as 'Instagram', msft.close 'Microsoft', orac.close 'Oracle', techy.close '騰訊'
6 from applistockprice aa
7 left join amznstockprice am
8 on aa.date= am.date
9 left join fbstockprice fb
10 on aa.date=fb.date
11 left join googlestockprice gg
12 on aa.date= gg.date
13 left join ibmstockprice ibm
14 on aa.date= ibm.date
15 left join intcstockprice ints
16 on aa.date= ints.date
17 left join msftstockprice msft
18 on aa.date= msft.date
19 left join oraclestockprice orac
20 on aa.date= orac.date
21 left join techystockprice techy
22 on aa.date=techy.date
23 where aa.date= ('2016-03-28');
24
25
```

Date	apple	amazon	Facebook	Google	IBM	Instagram	Microsoft	Oracle	騰訊
2016-03-28	105.190002	579.869995	113.690002	733.530029	148.399994	31.9	53.540001	40.619999	20.57

Select Statement #4

Select Statement #5

Select Statement #6

```
54
55 # To compare the close price between FB and TCENT in 2015
56 • select fb.date,fb.close facebook, tc.close techy
57 from fbstockprice as fb
58 left join techystockprice as tc
59 on fb.date= tc.date
60 where year(fb.date) = 2015;
61
```

100% 1:54

Result Grid Filter Rows: Search Export:

date	facebook	techy
2015-01-06	76.150002	15.290000
2015-01-07	76.150002	16.230000
2015-01-08	78.180000	16.600000
2015-01-09	77.739998	16.510000
2015-01-12	76.720001	16.200000
2015-01-13	76.449997	16.440001
2015-01-14	76.279999	16.170000
2015-01-15	74.050002	16.100000

Result 52

Read

Select Statement #7

```
1 use _weyda9345_project;
2
3 #7 select all the leaders' name and title from tencent
4 select CompanyCode,Name, title
5 from company_execs
6 where CompanyCode = 'Tcent';
7
8
9
```

CompanyCode	Name	title
1 TCENT	Mr. Chenye Xu	Co-Founder & Chief Information Officer
2 TCENT	Mr. Chi Ping Lau	Pres & Exec. Director
3 TCENT	Mr. Huateng Pony Ma	Co-Founder, Chairman & CEO
4 TCENT	Mr. Liqing Zeng	Co-Founder & Advisor Emeritus
5 TCENT	Mr. Zhidong Zhang	Co-Founder & Advisor Emeritus

Select Statement #8

```
BusTartarus-FL19.ad.colorado.edu - _weya9345_2exampractice
1 use _weya9345_project;
2
3
4 #8 connect all_profiles and company_execs
5 select * from company_execs
6 left join all_profiles
7 on company_execs.CompanyCode = all_profiles.CompanyCode;
```

Name	Title	Pay	Exercised	YearBorn	CompanyCode	CompanyCode	Address	City
Dr. John E. Kelly III	Exec. VP	1730936		1954	IBM	IBM	1 New Orchard Road	Armonk
Dr. Venkata S. Murthy Renduchintala	Chief Engineering Officer and Group Pre...	5253900		1965	INTEL	INTEL	2200 Mission College Boulevard	Santa Clara
Mr. Andrew R. Jassy	Chief Exec. Officer of Amazon Web Servi...	266232		1968	AMZN	AMZN	410 Terry Avenue North	Seattle
Mr. Bradford L. Smith	Pres & Chief Legal Officer	4327560		1959	MSFT	MSFT	One Microsoft Way	Redmond
Mr. Brian T. Olsavsky	Sr. VP & CFO	163200		1964	AMZN	AMZN	410 Terry Avenue North	Seattle
Mr. Cheney Xu	Co-Founder & Chief Information Officer			1971	TCENT	TCENT	Tencent Binhai Towers	Shenzhen
Mr. Chi Ping Lau	Pres & Exec. Director			1973	TCENT	TCENT	Tencent Binhai Towers	Shenzhen
Mr. Chris Kondo	Sr. Director of Corp. Accounting				APPL	APPL	One Apple Park Way	Cupertino
Mr. David C. Drummond J.D.	Sr. VP of Corp. Devel., Chief Legal Office...	669585	60167028	1963	GOOG	GOOG	1600 Amphitheatre Parkway	Mountain View
Mr. David M. Wehner	Chief Financial Officer	1262590		1969	FB	FB	1601 Willow Road	Menlo Park
Mr. Edward Screven	Exec. VP & Chief Corp. Architect	708619	20353500	1965	ORCL	ORCL	500 Oracle Parkway	Redwood Shores
Mr. Fletcher Previn	Chief Information Officer				IBM	IBM	1 New Orchard Road	Armonk
Mr. George S. Davis	Exec. VP & CFO			1958	INTEL	INTEL	2200 Mission College Boulevard	Santa Clara
Mr. Huateng Pony Ma	Co-Founder, Chairman & CEO			1972	TCENT	TCENT	Tencent Binhai Towers	Shenzhen
Mr. James J. Kavanaugh	Sr. VP & CFO	1609951		1967	IBM	IBM	1 New Orchard Road	Armonk
Mr. Jean-Philippe Courtois	Exec. VP and Pres of Microsoft Global S...	3775280		1961	MSFT	MSFT	One Microsoft Way	Redmond
Mr. Jeffrey A. Wilke	Chief Exec. Officer of Worldwide Consumer 255613			1967	AMZN	AMZN	410 Terry Avenue North	Seattle
Mr. Jeffrey E. Williams	Chief Operating Officer	5051818		1964	APPL	APPL	One Apple Park Way	Cupertino
Mr. Jeffrey M. Blackburn	Sr. VP of Bus. Devel.	178500		1970	AMZN	AMZN	410 Terry Avenue North	Seattle
Mr. Jeffrey O. Henley	Vice Chairman	658615	22751200	1945	ORCL	ORCL	500 Oracle Parkway	Redwood Shores
Mr. Jeffrey P. Bezos	Founder, Chairman, Pres & CEO	1681840		1964	AMZN	AMZN	410 Terry Avenue North	Seattle
Mr. Lawrence Edward Page	Co-Founder & Director			1973	GOOG	GOOG	1600 Amphitheatre Parkway	Mountain View
Mr. Lawrence J. Ellison	Co-Founder, Chairman & CTO	1662828	204707008	1944	ORCL	ORCL	500 Oracle Parkway	Redwood Shores
Mr. Liqing Zeng	Co-Founder & Advisor Emeritus			1970	TCENT	TCENT	Tencent Binhai Towers	Shenzhen
Mr. Luca Maestri	CFO & Sr. VP	5017804		1964	APPL	APPL	One Apple Park Way	Cupertino
Mr. Mark Elliot Zuckerberg	Founder, Chairman & CEO	22554543		1984	FB	FB	1601 Willow Road	Menlo Park
Mr. Martin J. Schroeter	Sr. VP of IBM Global Markets Global Fin...	2149949		1964	IBM	IBM	1 New Orchard Road	Armonk
Mr. Michael T. Schroeper	Chief Technology Officer	1333840	139010112	1975	FB	FB	1601 Willow Road	Menlo Park
Mr. Navin Shenoy	Exec. VP & GM of Data Platforms Group	2920300		1974	INTEL	INTEL	2200 Mission College Boulevard	Santa Clara
Mr. Robert H. Swan	CEO & Director	5808500		1960	INTEL	INTEL	2200 Mission College Boulevard	Santa Clara

Finished 2 queries with 45 records. 0.102 seconds.

Select Statement #9

```
BusTartarus-FL19.ad.colorado.edu - _weya9345_2exampractice
1 use _weya9345_project;
2
3 #9 select the CEOs' name, and yearborn from the Technology companies
4 select CompanyCode,Name, title, YearBorn
5 from company_execs
6 where CompanyCode in (
7   select CompanyCode from all_profiles where Sector = 'Technology')
8 Having title like '%CEO%';
9
10
```

CompanyCode	Name	title	YearBorn
APPL	Mr. Timothy D. Cook	CEO & Director	1961
IBM	Ms. Virginia M. Rometty	Chairman, Pres & CEO	1959
INTEL	Mr. Robert H. Swan	CEO & Director	1960
MSFT	Mr. Satya Nadella	CEO & Director	1967
ORCL	Ms. Safra Ada Catz	CEO & Director	1962

Select Statement #10

```
BusTartarus-FL19.ad.cc Results du - _weya9345_2exampractice
1 use _weya9345_project;
2
3 # These nine companies belong to which states? which state's companies provide the most job opportunities ?
4 select state,sum(Employees) 'total employees', count(companycode) 'numbers'from all_profiles
5 group by state
6 order by sum(Employees) desc;
7
```

state	total employees	numbers
WA	791500	2
CA	537526	5
NY	350600	1
	60860	1

Select Statement #11

```
3 #11: get the gap(max(high) - min(low)) of each company in each year;
4 CREATE VIEW `Apple` AS
5 select year(Date) 'year', max(high) - min(low) 'gap'
6 from aaplstockprice
7 group by year(Date);
8
9 CREATE VIEW `Amazon` AS
10 select year(Date) 'year', max(high) - min(low) 'gap'
11 from amznstockprice
12 group by year(Date);
13
14 CREATE VIEW `Facebook` AS
15 select year(Date) 'year', max(high) - min(low) 'gap'
16 from fbstockprice
17 group by year(Date);
18
19 CREATE VIEW `Google` AS
20 select year(Date) 'year', max(high) - min(low) 'gap'
21 from googlestockprice
22 group by year(Date);
23
24 CREATE VIEW `IBM` AS
25 select year(Date) 'year', max(high) - min(low) 'gap'
26 from ibmstockprice
27 group by year(Date);
28
29 CREATE VIEW `intel` AS
30 select year(Date) 'year', max(high) - min(low) 'gap'
31 from intcstockprice
32 group by year(Date);
33
34 CREATE VIEW `Microsoft` AS
35 select year(Date) 'year', max(high) - min(low) 'gap'
36 from msftstockprice
37 group by year(Date);
38
39 CREATE VIEW `Oracle_Corporation` AS
40 select year(Date) 'year', max(high) - min(low) 'gap'
41 from oraclestockprice
42 group by year(Date);
43
44 CREATE VIEW `TENCENT` AS
45 select year(Date) 'year', max(high) - min(low) 'gap'
46 from techystockprice
47 group by year(Date);
```

```

BusTartarus-FL19.ad.colorado.edu - _weya9345_exampractice
1 use _weya9345_project;
2 #11: get the gap(max(high) - min(low)) of each company in each year;
3 select ap.year, ap.gap 'gap of apple', az.gap 'gap of amazon', fb.gap 'gap of facebook', gl.gap 'gap of google',
4          ib.gap 'gap of IBM', it.gap 'gap of intel', ms.gap 'gap of microsoft',
5          oc.gap 'gap of oracle_corporation', tc.gap 'gap of tencent'
6 from apple ap
7 left join amazon az
8 on ap.year = az.year
9 left join facebook fb
10 on ap.year = fb.year
11 left join google gl
12 on ap.year = gl.year
13 left join ibm ib
14 on ap.year = ib.year
15 left join intel it
16 on ap.year = it.year
17 left join microsoft ms
18 on ap.year = ms.year
19 left join oracle_corporation oc
20 on ap.year = oc.year
21 left join tencent tc
22 on ap.year = tc.year

```

year	gap of apple	gap of amazon	gap of facebook	gap of google	gap of IBM	gap of intel	gap of microsoft	gap of oracle_corporation	gap of tencent
1 2000	4.397322	76.625	NULL	NULL	54.875	46.015625	39.15625	24.96875	NULL
2 2001	0.905893	16.865	NULL	NULL	40.949997	19.633751	16.637501	24.84	NULL
3 2002	0.915	15.97	NULL	NULL	72.380001	23.829999	14.605001	10.25	NULL
4 2003	0.877858	42.600003	NULL	NULL	21.370003	19.629998	7.450001	3.39	NULL
5 2004	3.456428	24.82	NULL	52.622753	18.529998	14.959999	6.190001	5.73	NULL
6 2005	6.308572	19.4	NULL	136.309075	27.25	6.950001	4.43	3.26	NULL
7 2006	6.142858	22.820002	NULL	90.3862	25.149994	9.879999	8.800001	7.69	NULL
8 2007	17.294286	64.789997	NULL	154.540756	32.690002	9.24	10.9	7.339999	NULL
9 2008	17.302857	62.75	NULL	224.194672	61.429993	14.28	18.459999	8.620001	NULL
10 2009	19.392856	98.280003	NULL	170.979126	51.090004	9.22	16.63	11.310001	NULL
11 2010	19.487142	79.849991	NULL	98.241775	31.529999	6.770001	8.85	11.03	1.692
12 2011	16.599999	86.120011	NULL	86.545608	48.259995	6.620001	5.809999	11.780001	2.224
13 2012	42.295719	92.109985	27.450001	108.523224	34.439987	10.04	6.690001	9.019998	3.272
14 2013	27.148571	159.880005	35.910002	211.945617	43.329987	5.940001	12.699999	8.479999	6.916
15 2014	49.242859	124.059998	30.32	124.484375	48.710007	14.400002	15.419998	11.27	5.599999
16 2015	42.539993	411.190002	38.650002	293.754914	44.650009	12.620001	17.129997	10.190003	7.88
17 2016	29.220001	373.210022	44.129997	153.619995	53.049995	10.680001	16.059997	8.869999	11.48
18 2017	62.439995	465.710022	68.739998	302.690002	43.659988	14.409999	25.549999	14.84	31.799999
19 2018	86.880005	879.98999	95.599998	303.78003	65.190003	15.559997	32.349998	11.079998	29.459999

Finished 2 queries with 20 records. 0.429 seconds.

Select Statement #12

```

BusTartarus-FL19.ad.colorado.edu - _weya9345_exampractice
1 use _weya9345_project;
2 #12:
3 # Robert Shiller, the famous economist from Yale, is arguing with Eugene Fama, another famous economist, regarding the stock market.
4 # Eugene believes the market is pretty efficient, so that returns on any given day are unrelated to any other day.
5 # Robert thinks Mondays are different: after the weekend people go into Wall Street on Monday feeling gloomy, and as a consequence returns should be lower.
6 # According to this, Is the stock average close price lowest on Monday?
7 drop view if exists 'avg_adj_close_price';
8 CREATE VIEW `avg_adj_close_price` AS
9 select dayname(ap.date) dayname,
10        avg(ap.Adj_Close) 'apple', avg(az.Adj_Close) 'amazon', avg(fb.Adj_Close) 'facebook', avg(gl.Adj_Close) 'google',
11        avg(ib.Adj_Close) 'IBM', avg(it.Adj_Close) 'intel', avg(ms.Adj_Close) 'microsoft',
12        avg(oc.Adj_Close) 'oracle_corporation', avg(tc.Adj_Close) 'tencent'
13 from aplstockprice ap
14 left join amznstockprice az
15 on ap.date = az.date
16 left join fbstockprice fb
17 on ap.date = fb.date
18 left join googlesockprice gl
19 on ap.date = gl.date
20 left join ibmstockprice ib
21 on ap.date = ib.date
22 left join intcstockprice it
23 on ap.date = it.date
24 left join msftstockprice ms
25 on ap.date = ms.date
26 left join oraclesockprice oc
27 on ap.date = oc.date
28 left join techstockprice tc
29 on ap.date = tc.date
30 group by dayname(ap.date);

```

