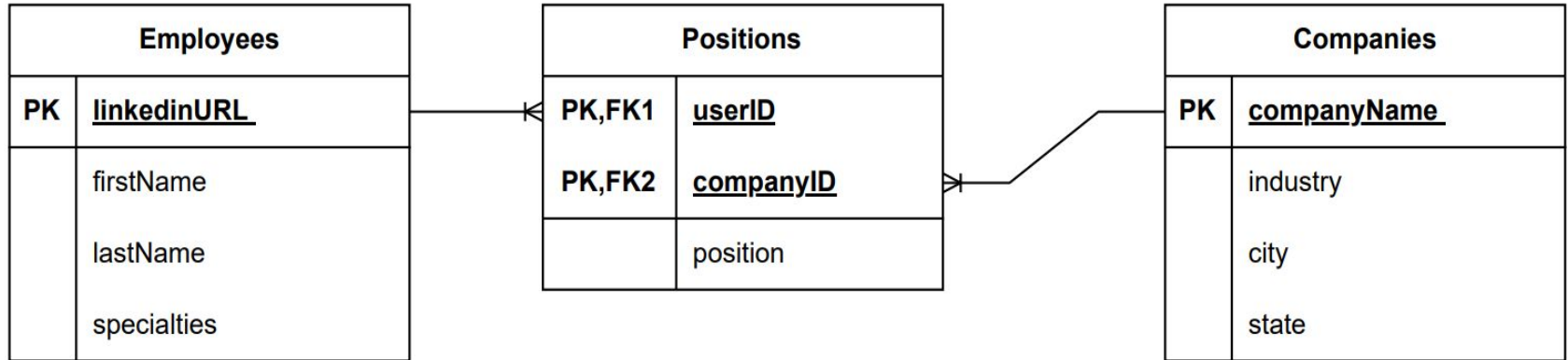# Database of Wall Street Company Data

By **Team 8**: Logan Johnston, Sam Ramdial, Ryan Shultz, Justin Khuu, Senaye Weldeberhan

Objectives

- Developing a better understanding of database design principles

- Completion of database according to Wall Street Company's guidelines

Database Logical Design

Normalization (BCNF) - Database Schema

Table: **Employees**
Columns:
        linkedinURL VARCHAR(255) NOT NULL
        firstName VARCHAR(255) NOT NULL
        lastName VARCHAR(255) NOT NULL
        specialties VARCHAR(255)
PRIMARY KEY (linkedinURL)

Table: **Companies**
Columns:
        companyName VARCHAR(255) NOT NULL
        industry VARCHAR(255) NOT NULL
        city VARCHAR(20) NOT NULL
        state CHAR(2) NOT NULL
PRIMARY KEY (companyName)

Table: **Positions**
Columns:
        userID VARCHAR(255) NOT NULL
        companyID VARCHAR(255) NOT NULL
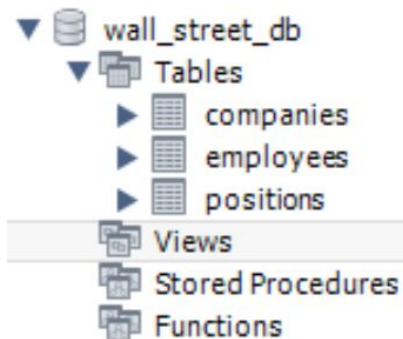   Position VARCHAR(255) NOT NULL
PRIMARY KEY (userID, companyID)
FOREIGN KEY (userID)
REFERENCES employees(linkedinURL)
FOREIGN KEY (companyID)
REFERENCES companies(companyName)

▼ 🛢 wall_street_db
    ▼ 🗂 Tables
        ▶ 🗂 companies
        ▶ 🗂 employees
        ▶ 🗂 positions
    🗂 Views
    🗂 Stored Procedures
    🗂 Functions

Security, Backup & Recovery

- Security
  - Update MySQL to get the most recent security patches
  - Make sure that all the data is encrypted so if there is a successful attack, the attackers wont gain much
  - Always monitor the database for any unusual activity
  - Implementing role based permissions can ensure that only the qualified person has access to certain information
- Backup & Recovery
  - Full backups are recommended weekly
  - Incremental backups are recommended daily to ensure any new data is backup
  - Export backups to an external hard drive, along with to the cloud
  - To restore data from the backups import the backup data into MySQL
  - Recovery will take as long as needed depending on the severity of the situation

Data Migration

1. **Gathering the Data:** Pulling information from CSV files using SQL import statements

2. **Organization:** Referring to the last weekly/daily backup, and using that to format the database properly

3. **Testing:** Making sure that the data listed functions properly within MySQL and software used to manage the database

```sql
LOAD DATA INFILE 'c:\\ProgramData\\MySQL\\MySQL Server 8.0\\Uploads\\wall_street_db.csv'
IGNORE
INTO TABLE employees
FIELDS TERMINATED BY ','
ENCLOSED BY '"'
LINES TERMINATED BY '\r\n'
IGNORE 1 ROWS
(@a,@b,@c,@d,@e,@f,@g,@h,@i)
set linkedinURL = @i, firstName = @a, lastName = @b, specialities = @e;

LOAD DATA INFILE 'c:\\ProgramData\\MySQL\\MySQL Server 8.0\\Uploads\\wall_street_db.csv'
IGNORE
INTO TABLE companies
FIELDS TERMINATED BY ','
ENCLOSED BY '"'
LINES TERMINATED BY '\r\n'
IGNORE 1 ROWS
(@a,@b,@c,@d,@e,@f,@g,@h,@i)
set companyName = @d, industry = @f, city = @g, state = @h;

LOAD DATA INFILE 'c:\\ProgramData\\MySQL\\MySQL Server 8.0\\Uploads\\wall_street_db.csv'
IGNORE
INTO TABLE positions
FIELDS TERMINATED BY ','
ENCLOSED BY '"'
LINES TERMINATED BY '\r\n'
IGNORE 1 ROWS
(@a,@b,@c,@d,@e,@f,@g,@h,@i)
set userID = @i, companyID = @d, position = @c;
```

MySQL Views

companyName	city	state
Baird	Chicago	IL
Bank of America	Chicago	IL
BMO Capital Markets	Chicago	IL
Goldman Sachs	Chicago	IL
Houlihan Lokey	Chicago	IL
KeyBanc Capital Markets	Chicago	IL
Lincoln International	Chicago	IL
William Blair	Chicago	IL

```sql
CREATE
    ALGORITHM = UNDEFINED
    DEFINER = `root`@`localhost`
    SQL SECURITY DEFINER
VIEW `employable_cities_view` AS
    SELECT
        `c`.`companyName` AS `companyName`,
        `c`.`city` AS `city`,
        `c`.`state` AS `state`
    FROM
        (`positions` `p`
        JOIN `companies` `c`)
    WHERE
        (`c`.`companyName` = `p`.`companyID`)
    GROUP BY `c`.`companyName`
    HAVING (COUNT(`p`.`userID`) > 20)
```

MySQL Views

```sql
CREATE
    ALGORITHM = UNDEFINED
    DEFINER = `root`@`localhost`
    SQL SECURITY DEFINER
VIEW `employee_information_view` AS
    SELECT
        `e`.`firstName` AS `firstName`,
        `e`.`lastName` AS `lastName`,
        `p`.`userID` AS `linkedinURL`,
        `p`.`companyID` AS `Company`
    FROM
        (`positions` `p`
        JOIN `employees` `e`)
    WHERE
        (`e`.`linkedinURL` = `p`.`userID`)
    ORDER BY `e`.`firstName`
```

firstName	lastName	linkedinURL	Company
Aaron	Newman	https://www.linkedin.com/in/aaron-newman-18...	DC Advisory
Aaron	Small	https://www.linkedin.com/in/aaron-small-a8487...	Macquarie Group
Aaron	Zwiebach	https://www.linkedin.com/in/aaron-zwiebach-1...	Evolution Credit Partners
Aaron	Solomon	https://www.linkedin.com/in/aaronj-solomon	BNY Mellon
Abhishek	Narang	https://www.linkedin.com/in/narang-abhishek	Goldman Sachs
Abhisheka	Jain	https://www.linkedin.com/in/abhisheka-jain-830...	Citi
Adam	Waite	https://www.linkedin.com/in/adam-waite-97726b8	D.A. Davidson
Adam	Bontrager	https://www.linkedin.com/in/adam-bontrager-7...	KeyBanc Capital Markets
Adam	Green	https://www.linkedin.com/in/adam-green-7007...	Livingstone Partners
Adam	F.	https://www.linkedin.com/in/adamcfine	Goldman Sachs
Aditi	Dekhane	https://www.linkedin.com/in/aditi-dekhane-551...	BNP Paribas CIB
Adnan	Q.	https://www.linkedin.com/in/adnan831	Kaufman Hall
AJ	Silverman	https://www.linkedin.com/in/ajsilverman1	Deloitte Investment Ban...
Al	Laufenberg	https://www.linkedin.com/in/al-laufenberg-4b3...	Stifel Nicolaus
Al	Torabi	https://www.linkedin.com/in/al-torabi-194a4a3a	Solomon Partners
Alden	Wong	https://www.linkedin.com/in/alden-wong-6980287	J.P. Morgan
Alec	Morgan	https://www.linkedin.com/in/alec-morgan-4866...	Houlihan Lokey

MySQL Views

```sql
CREATE
    ALGORITHM = UNDEFINED
    DEFINER = `root`@`localhost`
    SQL SECURITY DEFINER
VIEW `employee_numbers_view` AS
    SELECT
        `p`.`companyID` AS `Company`,
        COUNT(`p`.`userID`) AS `numEmployees`
    FROM
        ((`positions` `p`
        JOIN `companies` `c` ON ((`c`.`companyName` = `p`.`companyID`)))
        JOIN `employees` `e` ON ((`e`.`linkedinURL` = `p`.`userID`)))
    GROUP BY `Company`
```

	Company	numEmployees
	Apogem Capital	1
	AQ Technology Par...	1
	Arbor Investments	2
	Arlington Capital A...	1
	Armory Group	1
	Atlas Technology G...	1
	Auctus Capital Part...	1
	Avenel Financial Gr...	1
	B. Riley Financial	3
	Babson F.W. Olin ...	1
	Baird	28
	Baird	Kellogg MBA...	1
	Balmoral Advisors	2
	Bank of America	30
	Bank of America M...	6
	Bank of America M...	1
	Barclays	6

MySQL Stored Procedures

```sql
CREATE DEFINER=`root`@`localhost` PROCEDURE `GetEmployeeInfo`(IN url VARCHAR(255),
OUT employeeFirstName VARCHAR(255), OUT employeeLastName VARCHAR(255), OUT employeeCompanyName
VARCHAR(255), OUT employeePosition VARCHAR(255))
    READS SQL DATA
BEGIN
    SELECT firstName, lastName, companyName, position
    INTO employeeFirstName, employeeLastName, employeeCompanyName, employeePosition
    FROM full_table_view
    WHERE linkedinURL = url;
END
```

Challenges & Improvements

- Finding times that worked for us to get together, as well as establishing clear guidelines and deadlines for each timeline | Established a group chat where we could all communicate
- Figuring out how we wanted to break down the database when it came to normalization | Attended Office Hours and reached out to TA's for additional assistance
- Importing the data into MySQL - Error with the encoded file of the data | Saved the file with the proper encoding