# 1 HW #1: Installing Postgres

In order to complete the assignment, please install PostgreSQL on your computer and then verify that it is installed correctly by doing the following. Note that installing PostgreSQL can be difficult and I recommend researching it before beginning. The purpose of this homework is to load the four tables that we will be using during the rest of this semester into our database.

1. We first need to create a destination schema for our data. The command below will create a schema, "stocks2016." After running this, follow the commands in Figure B.1 to create tables within this schema.

   ```
   create schema stocks2016;
   commit;
   ```

2. Now that we have a destination for our data we need to load the following files into the database, all of which can be found on canvas. The commands can do this can be found in Figure B.2, but will need to be modified depending on the specific locations of your files.

3. Using the 2010 data, write a query that returns the following:[1]

   (a) All columns relating to CUSIP 45920010

   (b) All columns for the symbol "45920010" on the 7th of January

   (c) The difference between the bid and the ask for 45920010 on the 7th of January

   (d) The days when cusip 45920010 has a volume less than 5 million and a bid over $140.

4. We define the market capitalization as the number of shares outstanding multiplied by the price. Using this definition answer the following questions:

   (a) What is the market capitalization for permno 14593 on the first of February in 2010?

   (b) What are the permnos and market capitalizations for the companies with the top 5 market capitalization-days in 2010?

   (c) What are the permnos and market capitalizations for the companies with the top 5 market capitalizations on February 3rd 2010?

   (d) What are the permnos and market capitalizations for the companies with the bottom 5 market capitalization-days in 2010?

[1] If the query returns a significant number of rows, please only copy a few rows in your response.

   (e) What are the permnos and market capitalizations for the companies with the bottom 5 market capitalization-days on February 3rd 2010 with stocks that have a trading volume of less than 10 million?

5. We define the bid ask spread as the bid - ask. Generally this will be positive, but in our data set it is negative.[2]

   (a) Permno of the company with the smallest bid-ask spread in 2010 with a volume less than 25,000?

   (b) Permno of the company with the smallest bid-ask spread on February 8th 2010 that also had more than 500000 shares outstanding?

   (c) What is the permno and the bid-ask spread of the stock with the smallest bid-ask spread when the volume is less than 1,000?

6. The table fnd has information from company's annual reports. Using that dataset, write queries which answer the following questions.

   (a) Which company (ticker symbol) had the highest net income?

   (b) Which company (ticker symbol) had the highest net income in fiscal year 2011?

   (c) Which company (ticker symbol) had the lowest net income?

   (d) Which company (ticker symbol) had the lowest net income in fiscal year 2011?

   (e) Which company (ticker symbol), which had more than 1,000 employees, had the highest net income per employee in 2011?

   (f) Which company (ticker symbol), which had a net-income per employee over $1,000 had the largest number of employees? Keep units in mind!

7. Finally, make sure that you have a photo on your canvas page.

[2] See the data dictionary for more information.

```
create table stocks2016.d2010 (
        cusip varchar(8)
        , PERMNO bigint
        , PERMCO bigint
        , ISSUNO float
        , hsic float
        , retdate date
        , bid float
        , ask float
        , PRC float
        , VOL bigint
        , RET varchar(20)
        , SHROUT bigint);
commit;
create table stocks2016.d2011 (
        cusip varchar(8)
        , PERMNO bigint
        , PERMCO bigint
        , ISSUNO float
        , hsic float
        , retdate date
        , bid float
        , ask float
        , PRC float
        , VOL bigint
        , RET varchar(20
        , SHROUT bigint);
commit;
create table stocks2016.fnd (
        gvkey varchar(8)
        , datadate date
        , fyear int
        , indfmr varchar(4)
        , consol varchar(1)
        , popsrc varchar(1)
        , datafmt varchar(3)
        , tic varchar(8)
        , cusip varchar(11)
        , conm varchar(30)
        , fyr int
        , cash float
        , dp float
        , ebitda float
        , emp float
        , invt float
        , netinc float
        , ppent float
        , rev float
        , ui float
        , cik varchar(10)
            );
commit;
create table stocks2016.lnk (
        gvkey varchar(8)
        , linkprim varchar(1)
        , liid varchar(4)
        , linktype varchar(4)
        , lpermno bigint
        , lpermco bigint
        , usedflag int
        , linkdt date
        , linkenddt varchar(10)
        );
commit;
```

Figure B.1: Create Table Commands for the four tables in our database

```
COPY stocks2016.lnk FROM
        '/Users/ncross/Dropbox/USF/LoadingData/output/link.tdf'
        HEADER CSV DELIMITER E'\t';

COPY stocks2016.fnd FROM
        '/Users/ncross/Dropbox/USF/LoadingData/output/compm.tdf'
        HEADER CSV DELIMITER E'\t';

COPY stocks2016.d2010 FROM
        '/Users/ncross/Dropbox/USF/LoadingData/output/dsf2010.tdf'
        HEADER CSV DELIMITER E'\t';

 COPY stocks2016.d2011 FROM
         '/Users/ncross/Dropbox/USF/LoadingData/output/dsf2011.tdf'
        HEADER CSV DELIMITER E'\t';

commit;
```

Figure B.2: Copy Commands