There are several videos, this Is the first:

1. https://www.youtube.com/watch?v=zy6LZ3QRDu4&list=PLE7E8B7F4856C9B19&index=36
2. We have started.
3. We also opened terminal: mysql -u root -p but before enter and password.
4. We can actually drag and drop the sql file we downloaded from:

http://www.seanlahman.com/baseball-archive/statistics/

right into terminal.

So we end up with:

NCT:~ Tracksta6$ mysql -u root -p < /Users/Tracksta6/Dropbox/Computer\ Science/Java/Sabermetrics/lahman2016-sql/lahman2016.sql

Enter password:

NCT:~ Tracksta6$

1. There is a slight pause after password, and I believe that sql just actually imported the database.
2. Log in to mysql as normal:
   1. mysql –u root –p
   2. Password:
   3. mysql> SHOW databases;

+--------------------+

| Database |

+--------------------+

| information\_schema |

| lahman2016 |

| mysql |

| performance\_schema |

| student\_database |

| sys |

+--------------------+

1. 6 rows in set (0.01 sec)
2. So we just imported our database.
   1. mysql> USE lahman2016;
3. Note, there may be a chance that mysql can lock up, if you ask it to do something that is too complicated. Derek Banas said to just open up a new tab, log in:
   1. mysql –u root –p;
   2. mysql> SHOW processlist;

+----+------+-----------+------------+---------+------+----------+------------------+

| Id | User | Host | db | Command | Time | State | Info |

+----+------+-----------+------------+---------+------+----------+------------------+

| 16 | root | localhost | NULL | Sleep | 715 | | NULL |

| 18 | root | localhost | lahman2016 | Query | 0 | starting | SHOW processlist |

+----+------+-----------+------------+---------+------+----------+------------------+

1. 2 rows in set (0.00 sec)
2. If lahman2016 is blocking us up, enter:
3. mysql > KILL query 16;
4. That was a side note.
5. Use the database and view the tables inside:

mysql> SHOW tables;

+----------------------+

| Tables\_in\_lahman2016 |

+----------------------+

| AllstarFull |

| Appearances |

| AwardsManagers |

| AwardsPlayers |

| AwardsShareManagers |

| AwardsSharePlayers |

| Batting |

| BattingPost |

| CollegePlaying |

| Fielding |

| FieldingOF |

| FieldingOFsplit |

| FieldingPost |

| HallOfFame |

| HomeGames |

| Managers |

| ManagersHalf |

| Master |

| Parks |

| Pitching |

| PitchingPost |

| Salaries |

| Schools |

| SeriesPost |

| Teams |

| TeamsFranchises |

| TeamsHalf |

+----------------------+

1. 27 rows in set (0.00 sec)
2. We are just going to be using Salaries and Teams to view what the teams paid.

mysql> SELECT s.yearID, s.teamID, SUM(s.salary) AS SALARY

FROM salaries s, teams t

WHERE s.yearID = 2002

AND s.teamID = t.teamID

AND s.yearID = t.yearID

GROUP BY s.teamID;

+--------+--------+-----------+

| yearID | teamID | SALARY |

+--------+--------+-----------+

| 2002 | ANA | 61721667 |

| 2002 | ARI | 102819999 |

| 2002 | ATL | 92870367 |

| 2002 | BAL | 60493487 |

| 2002 | BOS | 108366060 |

| 2002 | CHA | 57052833 |

| 2002 | CHN | 75690833 |

| 2002 | CIN | 45050390 |

| 2002 | CLE | 78909449 |

| 2002 | COL | 56851043 |

| 2002 | DET | 55048000 |

| 2002 | FLO | 41979917 |

| 2002 | HOU | 63448417 |

| 2002 | KCA | 47257000 |

| 2002 | LAN | 94850953 |

| 2002 | MIL | 50287833 |

| 2002 | MIN | 40425000 |

| 2002 | MON | 38670500 |

| 2002 | NYA | 125928583 |

| 2002 | NYN | 94633593 |

| 2002 | OAK | 40004167 |

| 2002 | PHI | 57954999 |

| 2002 | PIT | 42323599 |

| 2002 | SDN | 41425000 |

| 2002 | SEA | 80282668 |

| 2002 | SFN | 78299835 |

| 2002 | SLN | 74660875 |

| 2002 | TBA | 34380000 |

| 2002 | TEX | 105526122 |

| 2002 | TOR | 76864333 |

+--------+--------+-----------+

mysql> SELECT s.yearID, s.teamID, SUM(s.salary) AS SALARY

FROM salaries s, teams t

WHERE s.yearID = 2002

AND s.teamID = t.teamID

AND s.yearID = t.yearID

GROUP BY s.teamID

ORDER BY SALARY DESC;

+--------+--------+-----------+

| yearID | teamID | SALARY |

+--------+--------+-----------+

| 2002 | NYA | 125928583 |

| 2002 | BOS | 108366060 |

| 2002 | TEX | 105526122 |

| 2002 | ARI | 102819999 |

| 2002 | LAN | 94850953 |

| 2002 | NYN | 94633593 |

| 2002 | ATL | 92870367 |

| 2002 | SEA | 80282668 |

| 2002 | CLE | 78909449 |

| 2002 | SFN | 78299835 |

| 2002 | TOR | 76864333 |

| 2002 | CHN | 75690833 |

| 2002 | SLN | 74660875 |

| 2002 | HOU | 63448417 |

| 2002 | ANA | 61721667 |

| 2002 | BAL | 60493487 |

| 2002 | PHI | 57954999 |

| 2002 | CHA | 57052833 |

| 2002 | COL | 56851043 |

| 2002 | DET | 55048000 |

| 2002 | MIL | 50287833 |

| 2002 | KCA | 47257000 |

| 2002 | CIN | 45050390 |

| 2002 | PIT | 42323599 |

| 2002 | FLO | 41979917 |

| 2002 | SDN | 41425000 |

| 2002 | MIN | 40425000 |

| 2002 | OAK | 40004167 |

| 2002 | MON | 38670500 |

| 2002 | TBA | 34380000 |

+--------+--------+-----------+

30 rows in set (0.03 sec)

mysql> SELECT t.W AS WINS FROM teams t Where t.yearID = 2002;

+------+

| WINS |

+------+

| 99 |

| 98 |

| 101 |

| 67 |

| 93 |

| 81 |

| 67 |

| 78 |

| 74 |

| 73 |

| 55 |

| 79 |

| 84 |

| 62 |

| 92 |

| 56 |

| 94 |

| 83 |

| 103 |

| 75 |

| 103 |

| 80 |

| 72 |

| 66 |

| 93 |

| 95 |

| 97 |

| 55 |

| 72 |

| 78 |

+------+

30 rows in set (0.06 sec)

1. Note – I am having trouble getting the exact same example that he is showing. We are using the same basic everything, but the SQL version that downloaded was updated from the one he downloaded. I think that there are some slight differences that don’t allow for an exact copy on everything.

1. mysql> SELECT b.playerID,

-> CONCAT(m.nameFirst, " ", m.nameLast) AS NAME,

-> SUM(b.RBI) AS RBI

-> FROM Batting b, Master m

-> WHERE b.playerID = m.playerID

-> GROUP BY playerID

-> ORDER BY RBI DESC

-> LIMIT 50;

1. THIS DID NOT WORK.
2. I could not get this to work, since I am pulling in the playerID and the names and therefore need to group by both. Not sure why.
3. NOTE: This is a decent size database for my laptop. It took about 4 minutes to sort through, and make the list below.

mysql> SELECT b.playerID,

CONCAT(m.nameFirst, " ", m.nameLast) AS NAME,

SUM(b.RBI) AS RBI

FROM Batting b, Master m

Where b.playerID = m.playerID

GROUP BY b.playerID, NAME

ORDER BY RBI DESC

LIMIT 50;

+-----------+------------------+------+

| playerID | NAME | RBI |

+-----------+------------------+------+

| aaronha01 | Hank Aaron | 2297 |

| ruthba01 | Babe Ruth | 2217 |

| rodrial01 | Alex Rodriguez | 2086 |

| ansonca01 | Cap Anson | 2076 |

| bondsba01 | Barry Bonds | 1996 |

| gehrilo01 | Lou Gehrig | 1995 |

| musiast01 | Stan Musial | 1951 |

| cobbty01 | Ty Cobb | 1937 |

| foxxji01 | Jimmie Foxx | 1922 |

| murraed02 | Eddie Murray | 1917 |

| mayswi01 | Willie Mays | 1903 |

| ottme01 | Mel Ott | 1860 |

| yastrca01 | Carl Yastrzemski | 1844 |

| willite01 | Ted Williams | 1839 |

| griffke02 | Ken Griffey | 1836 |

| palmera01 | Rafael Palmeiro | 1835 |

| winfida01 | Dave Winfield | 1833 |

| ramirma02 | Manny Ramirez | 1831 |

| simmoal01 | Al Simmons | 1827 |

| pujolal01 | Albert Pujols | 1817 |

| robinfr02 | Frank Robinson | 1812 |

| ortizda01 | David Ortiz | 1768 |

| wagneho01 | Honus Wagner | 1732 |

| thomafr04 | Frank Thomas | 1704 |

| jacksre01 | Reggie Jackson | 1702 |

| thomeji01 | Jim Thome | 1699 |

| ripkeca01 | Cal Ripken | 1695 |

| sheffga01 | Gary Sheffield | 1676 |

| sosasa01 | Sammy Sosa | 1667 |

| perezto01 | Tony Perez | 1652 |

| bankser01 | Ernie Banks | 1636 |

| baineha01 | Harold Baines | 1628 |

| jonesch06 | Chipper Jones | 1623 |

| gosligo01 | Goose Goslin | 1609 |

| lajoina01 | Nap Lajoie | 1599 |

| brettge01 | George Brett | 1595 |

| schmimi01 | Mike Schmidt | 1595 |

| dawsoan01 | Andre Dawson | 1591 |

| hornsro01 | Rogers Hornsby | 1584 |

| killeha01 | Harmon Killebrew | 1584 |

| kalinal01 | Al Kaline | 1583 |

| becklja01 | Jake Beckley | 1575 |

| beltrad01 | Adrian Beltre | 1571 |

| mccovwi01 | Willie McCovey | 1555 |

| cabremi01 | Miguel Cabrera | 1553 |

| mcgrifr01 | Fred McGriff | 1550 |

| stargwi01 | Willie Stargell | 1540 |

| heilmha01 | Harry Heilmann | 1539 |

| dimagjo01 | Joe DiMaggio | 1537 |

| beltrca01 | Carlos Beltran | 1536 |

+-----------+------------------+------+

50 rows in set (3 min 27.92 sec)

* The next 2 are about the same, using RBI’s and HR’s
* I have run limits from 50 down to 5. They all take about 5 minutes.

mysql> SELECT b.playerID,

CONCAT(m.nameFirst, " ", m.nameLast) AS NAME,

SUM(b.RBI) AS RBI FROM Batting b, Master m

Where b.playerID = m.playerID

GROUP BY b.playerID, NAME

ORDER BY RBI DESC

LIMIT 5;

mysql> SELECT b.playerID, CONCAT(m.nameFirst, " ", m.nameLast) AS NAME, SUM(b.RBI) AS RBI FROM Batting b, Master m Where b.playerID = m.playerID GROUP BY b.playerID, NAME ORDER BY RBI DESC LIMIT 5;

+-----------+----------------+------+

| playerID | NAME | RBI |

+-----------+----------------+------+

| aaronha01 | Hank Aaron | 2297 |

| ruthba01 | Babe Ruth | 2217 |

| rodrial01 | Alex Rodriguez | 2086 |

| ansonca01 | Cap Anson | 2076 |

| bondsba01 | Barry Bonds | 1996 |

+-----------+----------------+------+

5 rows in set (3 min 26.25 sec)

mysql> SELECT b.playerID,

-> CONCAT(m.nameFirst, " ", m.nameLast) AS NAME,

-> SUM(b.HR) AS HR

-> FROM Batting b, Master m

-> Where b.playerID = m.playerID

-> GROUP BY b.playerID, NAME

-> ORDER BY HR DESC

-> LIMIT 10;

mysql> SELECT b.playerID, CONCAT(m.nameFirst, " ", m.nameLast) AS NAME, SUM(b.HR) AS HR FROM Batting b, Master m Where b.playerID = m.playerID GROUP BY b.playerID, NAME ORDER BY HR DESC LIMIT 10;

+-----------+----------------+------+

| playerID | NAME | HR |

+-----------+----------------+------+

| bondsba01 | Barry Bonds | 762 |

| aaronha01 | Hank Aaron | 755 |

| ruthba01 | Babe Ruth | 714 |

| rodrial01 | Alex Rodriguez | 696 |

| mayswi01 | Willie Mays | 660 |

| griffke02 | Ken Griffey | 630 |

| thomeji01 | Jim Thome | 612 |

| sosasa01 | Sammy Sosa | 609 |

| pujolal01 | Albert Pujols | 591 |

| robinfr02 | Frank Robinson | 586 |

+-----------+----------------+------+

10 rows in set (3 min 48.09 sec)

mysql> SELECT b.playerID,

-> CONCAT(m.nameFirst, " ", m.nameLast) AS NAME,

-> SUM(b.AB) AS AB,

-> (SUM(b.h)/SUM(b.AB)) AS AVG

-> FROM Batting b, Master m

-> Where b.playerID = m.playerID

-> GROUP BY b.playerID, NAME

-> HAVING AB > 1000

-> ORDER BY AVG DESC

-> LIMIT 5;

+-----------+----------------------+-------+--------+

| playerID | NAME | AB | AVG |

+-----------+----------------------+-------+--------+

| cobbty01 | Ty Cobb | 11434 | 0.3664 |

| barnero01 | Ross Barnes | 2392 | 0.3591 |

| hornsro01 | Rogers Hornsby | 8173 | 0.3585 |

| jacksjo01 | Shoeless Joe Jackson | 4981 | 0.3558 |

| meyerle01 | Levi Meyerle | 1443 | 0.3555 |

+-----------+----------------------+-------+--------+

5 rows in set (3 min 24.36 sec)