



# Lecture 2

---

## Tables

Slides created by John DeNero ([denero@berkeley.edu](mailto:denero@berkeley.edu)) and Ani Adhikari ([adhikari@berkeley.edu](mailto:adhikari@berkeley.edu))

# **Announcements**

# Python

# Programming Languages

---

- Python is popular both for data science & general software development
- Mastering the language fundamentals is critical
- Learn through practice, not by reading or listening
- Follow along: [datahub.berkeley.edu](https://datahub.berkeley.edu)

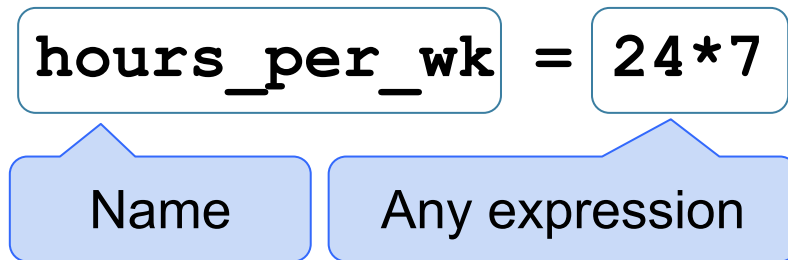
(Demo)

---

**Names**

# Assignment Statements

---



- Statements don't have a value; they perform an action
- An assignment statement changes the meaning of the name to the left of the = symbol
- The name is bound to a value (not an equation)

(Demo)

---

# Call Expressions

# Anatomy of a Call Expression

---

What  
function  
to call

Argument to the  
function

f(27)

"Call f on 27."

---



# Anatomy of a Call Expression

---

What  
function  
to call

First argument

Second  
argument

`max` (`15`, `27`)

(Demo)

---

# Tables

# Table Structure

---

- A Table is a sequence of labeled columns
- Each row represents one individual
- Data within a column represents one attribute of the individuals

The diagram illustrates a table structure with three columns: Name, Code, and Area (m2). The first row contains 'California', 'CA', and '163696'. The second row contains 'Nevada', 'NV', and '110567'. Annotations include a 'Label' bubble pointing to the 'Code' header, a 'Row' bubble pointing to the 'Nevada' cell, and a 'Column' bubble pointing to the 'NV' cell. A blue box highlights the 'NV' cell, and a blue line highlights the 'Nevada' row. The text '(Demo)' is located at the bottom right of the table area.

Name	Code	Area (m2)
California	CA	163696
Nevada	NV	110567

(Demo)

# Some Table Operations

---

- `t.select(label)` - constructs a new table with just the specified columns
  - `t.drop(label)` - constructs a new table in which the specified columns are omitted
  - `t.sort(label)` - constructs a new table with rows sorted by the specified column
  - `t.where(label, condition)` - constructs a new table with just the rows that match the condition
-

# Discussion Question

---

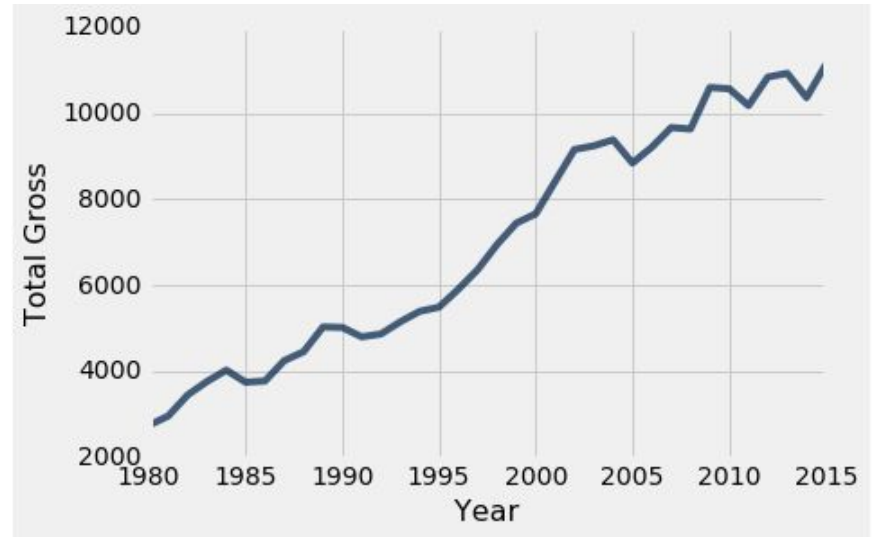
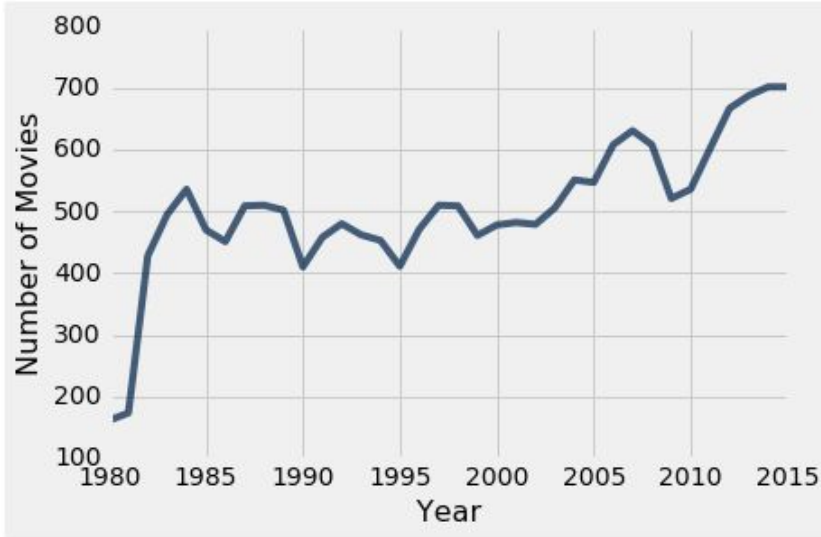
`nba` table:

How to display just the row corresponding to the player who had the highest salary?

(Demo)

---

# Discussion Question



Compare the plots over the years 2008-2010. Try to explain what you see, in relation to major events of 2008.