Lecture 6

Census
Announcements
Table Review
# Table Structure

- A Table is a sequence of labeled columns
- Labels are strings
- Columns are arrays, all with the same length

<table>
<thead>
<tr>
<th>Name</th>
<th>Code</th>
<th>Area (m²)</th>
</tr>
</thead>
<tbody>
<tr>
<td>California</td>
<td>CA</td>
<td>163696</td>
</tr>
<tr>
<td>Nevada</td>
<td>NV</td>
<td>110567</td>
</tr>
</tbody>
</table>
Table Methods

- Creating and extending tables:
  - \texttt{Table().with\_column} and \texttt{Table.read\_table}
- Finding the size: \texttt{num\_rows} and \texttt{num\_columns}
- Referring to columns: labels, relabeling, and indices
  - \texttt{labels} and \texttt{relabeled}; column indices start at 0
- Accessing data in a column
  - \texttt{column} takes a label or index and returns an array
- Using array methods to work with data in columns
  - \texttt{item, sum, min, max}, and so on
- Creating new tables containing some of the original columns:
  - \texttt{select, drop}

(Demo)
Manipulating Rows

- `t.sort(column)` sorts the rows in increasing order
- `t.take(row_numbers)` keeps the numbered rows
  - Each row has an index, starting at 0
- `t.where(column, are.condition)` keeps all rows for which a column's value satisfies a condition
- `t.where(column, value)` keeps all rows for which a column's value equals some particular value
- `t.with_row` makes a new table that has another row
Lists
Lists are Generic Sequences

A list is a sequence of values (just like an array), but the values can all have different types

\[ [2+3, \ 'four', \ \text{Table().with\_column('K', \ [3,\ 4])}] \]

- Lists can be used to create table rows.
- If you create a table column from a list, it will be converted to an array automatically

(Demo)
The table `nba` has columns `NAME`, `POSITION`, and `SALARY`.

a) Create an array containing the names of all point guards (PG) who make more than $15M/year

```javascript
nba.where(1, 'PG').where(2, are.above(15)).column(0)
```

b) After evaluating these two expressions in order, what's the result of the second one?

```javascript
nba.with_row([['Samosa', 'Mascot', 100]])
nba.where('NAME', are.containing('Samo'))
```
Census Data
The Decennial Census

- Every ten years, the Census Bureau counts how many people there are in the U.S.
- In between censuses, the Bureau estimates how many people there are each year.
- Article 1, Section 2 of the Constitution:
  - “Representatives and direct Taxes shall be apportioned among the several States … according to their respective Numbers …”
Analyzing Census Data

Leads to the discovery of interesting features and trends in the population

(Demo)
Census Table Description

- Values have column-dependent interpretations
  - The SEX column: 1 is Male, 2 is Female
  - The POPESTIMATE2010 column: 7/1/2010 estimate

- In this table, some rows are sums of other rows
  - The SEX column: 0 is Total (of Male + Female)
  - The AGE column: 999 is Total of all ages

- Numeric codes are often used for storage efficiency

- Values in a column have the same type, but are not necessarily comparable (AGE 12 vs AGE 999)

Growth Rate

- Growth rate = $g$ (for example 3%, or 0.03)
- Initial value $x$, final value $y$ after $t$ periods of time

Value after 1 period = $x + xg = x \times (1+g)$

Value after 2 periods = $x(1+g)(1+g) = x \times (1+g)^2$

Value after $t$ periods = $y = x \times (1+g)^t$

So $(1+g)^t = \frac{y}{x}$ and so $1+g = \left(\frac{y}{x}\right)^{\frac{1}{t}}$

So $g = \left(\frac{y}{x}\right)^{\frac{1}{t}} - 1$