Activity-based models

- At least 30 models established or under way in US
- Demand models trending to the disaggregate
- Big data
- Eg. PSRC (Seattle)
  - 1.5 million parcels
  - 3.6 million people
  - 11 million trips
  - 14 million activities
Terminology

Visualization
A technique for creating images, diagrams or animations to communicate a message

Analytics
The discovery and communication of meaningful patterns in data
Terminology

IPython Notebook

An interactive computational environment, in which you can combine code execution, rich text, math, plots and rich media

- For interactive authoring of literate computations
- Open source (BSD)
# Mode use by district

- **Mode Use by District**
- **Title**: Mode use by district, sublabel: "Trip", standard: true, legend: "top-right", title: "Trip", yscale: "linear", subtitle: "City", palette: "viridis"
- **Data**: bar, legend: "false"

## Activity sequence

- **Time**
- **Person Number**

<table>
<thead>
<tr>
<th>Person Number</th>
<th>Home</th>
<th>Work</th>
<th>School</th>
<th>Escort</th>
<th>Pers Bus</th>
<th>Home</th>
<th>Pers Bus</th>
</tr>
</thead>
<tbody>
<tr>
<td>1996086</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>1996089</td>
<td>2</td>
<td>2</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>1996090</td>
<td>2</td>
<td>2</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>2</td>
<td>2</td>
</tr>
</tbody>
</table>

## Mode by income and time

- **Mode by Income and Time**
- **Code**

```
In [3]: from matplotlib.display import display
import activity_data
import pandas
import numpy
df = activity_data.activity_table()

In [4]: df['income'] = pandas.qcut(df['income'], 4, labels=['A', 'B', 'C', 'D'])
plot(df['income'], df['time'], kind='line', legend='false', title='Mode by Income and Time', xlabel='Time', ylabel='Mode', color=['red', 'blue', 'green', 'purple'])
```
Awesome free tools!

- IPython Notebook for interactive computing
- Pandas for data wrangling, queries, pivots, joins and more
- Matplotlib, Bokeh, Seaborn etc… for beautiful charts
- “Data Scientists” create notebooks or apps for analysts
- But… what about mapping?
  - Existing libraries fell over after a few thousand geometries
  - A real need for something better
Terminology

Visual Analytics
Focuses on analytical reasoning facilitated by interactive visual interfaces

Big Data
Data sets so large or complex that traditional data processing applications are inadequate
Visualizations and disaggregate models

- Disaggregate spatial data is noisy and unwieldy
- Aggregations are lossy
- File-based workflows are slow with xGB datasets
- Creativity is very hard with interruptions
- Traditional tools?
- Need better support for visual information seeking
  - *Overview first, zoom and filter, then details-on demand*
What if we could...

- Keep data disaggregate
- Have instant access to data (in core memory)
- Map in 3D
- Animate time and explore space with high fps
- Filter, summarize and query on demand
- Graph, chart and prepare analytics in an open interactive computing system
Demonstration

Numbers to keep in mind
- 11 million trips
- 14 million activities
- 3.6 million people
- 1.5 million parcels
- 40,000 links

- Load time (<1 min)
- Query and filter time (<1 sec)
- Frame rate (target > 60 fps)
- RAM (up to 6GB)
- Mainstream hardware

*Soundcast ABM data courtesy PSRC*
Demonstration
Conclusions

A new tool for everyone
- 3D mapping and animations
- Visual analytics
- Large-scale spatial data

Interactive computing with IPython
- De facto data and science platform

Let’s work together!
- ABM applications
- Other large data sets

Daniel Florian
dan@inrosoftware.com
ABM Visualization and Analytics
Challenges and Solutions

Thanks!

Daniel Florian
dan@inrosoftware.com