Data Analysis and Visualization with Python

Avi Persin
avraham.persin@nasa.gov
Room 520
212-678-5633
Goals

1. **GISTEMP**
   - What / Why?
   - ccc-gistemp
   - New Interface
   - GISTEMP Plotter

2. **Station Locator**
   - Leaflet Map
   - GeoJson
   - Marker Clustering

3. **Global Basemaps**
   - Matplotlib Basemaps

4. **3D Graphics**
   - Mayavi
   - Climate Model
What / Why?

What:
- Analysis of NOAA (GHCN), SCAR, Antarctica
- In-house graphics plotting library

Why:
- Heterogeneous code base (Fortran, .sh, .exe)
- Multiple commands to execute
- Data validation
- Difficult to maintain and extend
CCC-GISTEMP

Features:
- < Python 2.7
- No Graphics
- No Optimizations
- Binary Reads
New GISTEMP

Features:

- One stop shop
- Python 3.x support
- NumPy
- Compressed .npz files
- Alter variables
- Data Visualizations
- Matplotlib to D3
Matplotlib and D3

- SVG vs Canvas
- Tex Support via MathJax
- Real time data retrieval via Ajax
- Custom interactive plugins
Moving Forward

- Export as html/js file
- Additional smoothing techniques
- Allow for additional dataset/input file types
- Pick format for table data files
Modules and Libraries

**Python:**
- Anaconda
- Matplotlib
- NumPy
- SciPy
- Mayavi
- mpld3
- Pandas
- IPython Notebook

**JavaScript:**
- D3.js
- Leaflet.js
- GeoJSON