## Grammar of Graphics

datascience@berkeley

### Graphics Production Language (GPL)

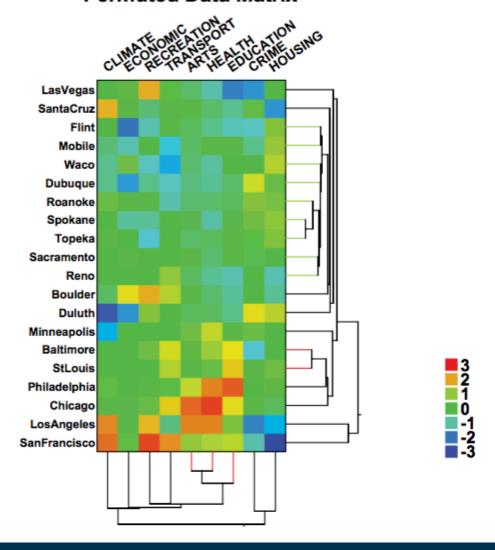
```
PAGE: begin
 SCALE: time(dim(1))
 SCALE: interval(dim(2), min(0), max(1.0d8), unit(currency.dollar))
 SCALE: cat(dim(3))
 SCALE: cat(dim(4))
 COORD: rect(dim(3,4), rect(dim(1,2)))
 GUIDE: axis(dim(1), label("Date"), format("mm/dd/yy"))
 GUIDE: axis(dim(2), label("Revenue"))
 GUIDE: axis(dim(3), label("Company"))
 GUIDE: axis(dim(4), label("State"))
 GUIDE: legend(aesthetic(color.blue), dim(1))
 ELEMENT: point(position(d*r*c*s), color(pe))
PAGE: end
```

#### Graphics Production Language (GPL)

```
DATA: unode = col(source("col tree"))
DATA: vnode = col(source("col tree"))
DATA: xnode = col(source("row tree"))
DATA: ynode = col(source("row tree"))
DATA: x = reshape.rect(climate, economic, recreation, transport,
         arts, health, education, crime, housing, "colname")
DATA: y = reshape.rect(climate, economic, recreation, transport, arts,
         health, education, crime, housing, "rowname")
DATA: d = reshape.rect(climate, economic, recreation, transport,
         arts, health, education, crime, housing, "value")
GRAPH: begin(origin(0, 0), scale(4cm, 8cm))
 ELEMENT: polygon(position(bin.rect(x*y)), color.hue(d))
GRAPH: end
GRAPH: begin(origin(0, -2cm), scale(4cm, 2cm))
 COORD: transpose(dim(1, 2), reflect(dim(2)))
 ELEMENT: edge(position(link.join(unode*vnode)))
GRAPH: end
GRAPH: begin(origin(4cm, 0cm), scale(2cm, 8cm))
 ELEMENT: edge(position(link.join(xnode*ynode)))
GRAPH: end
```

### Graphics Production Language (GPL)

#### **Permuted Data Matrix**

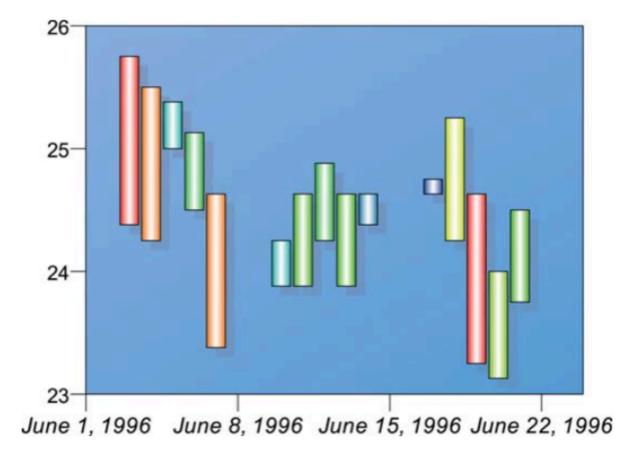


#### Visualization Markup Language (VizML)

XML-based (like SVG)

#### Visualization Markup Language (VizML)

#### XML-based (like SVG)



#### Visualization Markup Language (VizML)

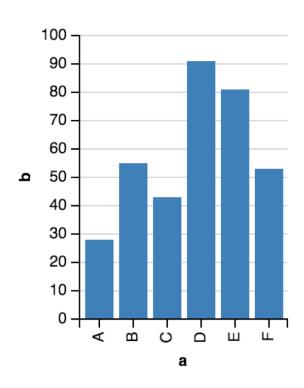
```
<?xml version="1.0"?>
<qraph>
  <source id="data">
      <fileAccess fileName="stockValues.csv" separator=","
header="true"/>
  </source>
  <variable id="date" sourceName="date" source="data" cate-</pre>
gorical="false">
      <meta>
          <dateTimeFormat>
             <dateFormat>
                 <dayOfMonth/>
                 <monthName long="true"/>
                 <year showCentury="true"/>
             </dateFormat>
          </dateTimeFormat>
      </meta>
   </variable>
  <variable id="high" sourceName="high" source="data" cate-</pre>
gorical="false"/>
  <variable id="low" sourceName="low" source="data" cate-</pre>
gorical="false"/>
  <variable id="close" sourceName="close" source="data"</pre>
             categorical="false"/>
   <variable id="range" categorical="false" expres-</pre>
sion="high+low"
             source="data"/>
  <domain categorical="false" id="dateDomain">
      <interval min="1996-06-02" max="1996-06-23"/>
  </domain>
```

### **VizQL**

Proprietary visualization grammar underlying Tableau

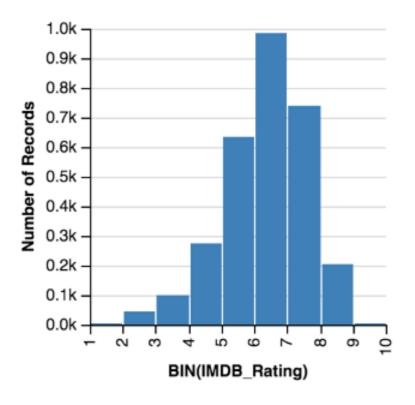
Evolved from Polaris system at Stanford (extends pivot tables)

#### Vega and Vega-Lite



## Vega-Lite

```
"data": {"url": "data/movies.json"},
"mark": "bar",
"encoding": {
    "x": {
        "bin": {"maxbins": 10},
        "field": "IMDB_Rating",
        "type": "quantitative"
    },
    "y": {
        "aggregate": "count",
        "field": "*",
        "type": "quantitative"
    }
}
```



## Vega-Lite

```
"data": {"url": "data/movies.json"},
                                                   110 -
"mark": "point",
                                                                                                    Number of Records
"encoding": {
                                                   100 -
                                                                                                        50
                                               BIN(Rotten_Tomatoes_Rating)
  "x": {
                                                    90 -
                                                                                                         100
    "bin": {"maxbins": 10},
                                                                                             0
                                                    80 -
    "field": "IMDB_Rating",
                                                                                                         150
                                                                           0
    "type": "quantitative"
                                                    70 -
  },
                                                    60 -
                                                    50 -
    "bin": {"maxbins": 10},
                                                    40 -
    "field": "Rotten_Tomatoes_Rating",
    "type": "quantitative"
                                                    30 -
  },
                                                    20 -
  "size": {
                                                    10 -
    "aggregate": "count",
    "field": "*",
                                                     0
    "type": "quantitative"
                                                           ^{\circ}
                                                                3
                                                                  BIN(IMDB_Rating)
```

# Berkeley school of information