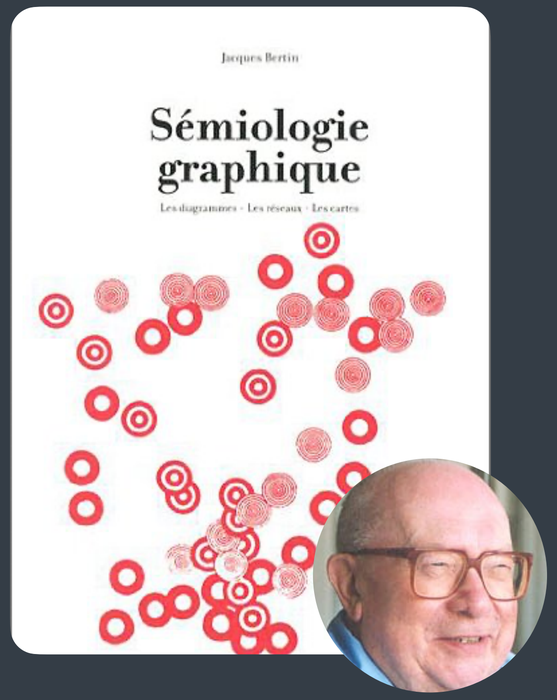
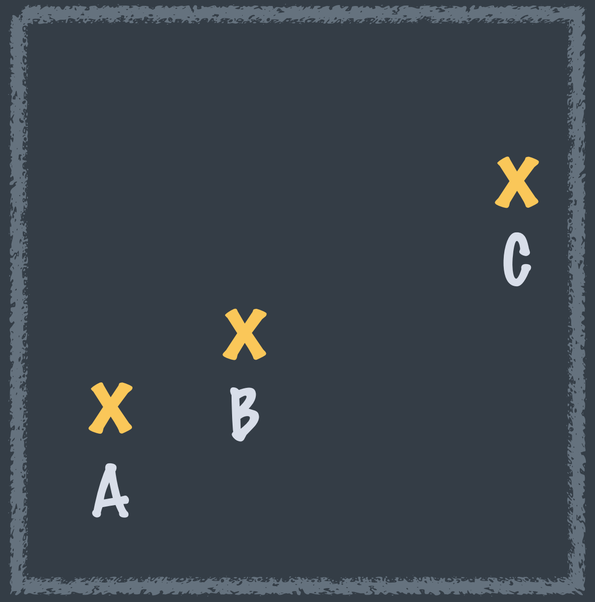
Towards *Effective* Interaction
With Data Visualization

Arvind Satyanarayan
@arvindsatya1

MIT Visualization Group
@mitvis • vis.csail.mit.edu

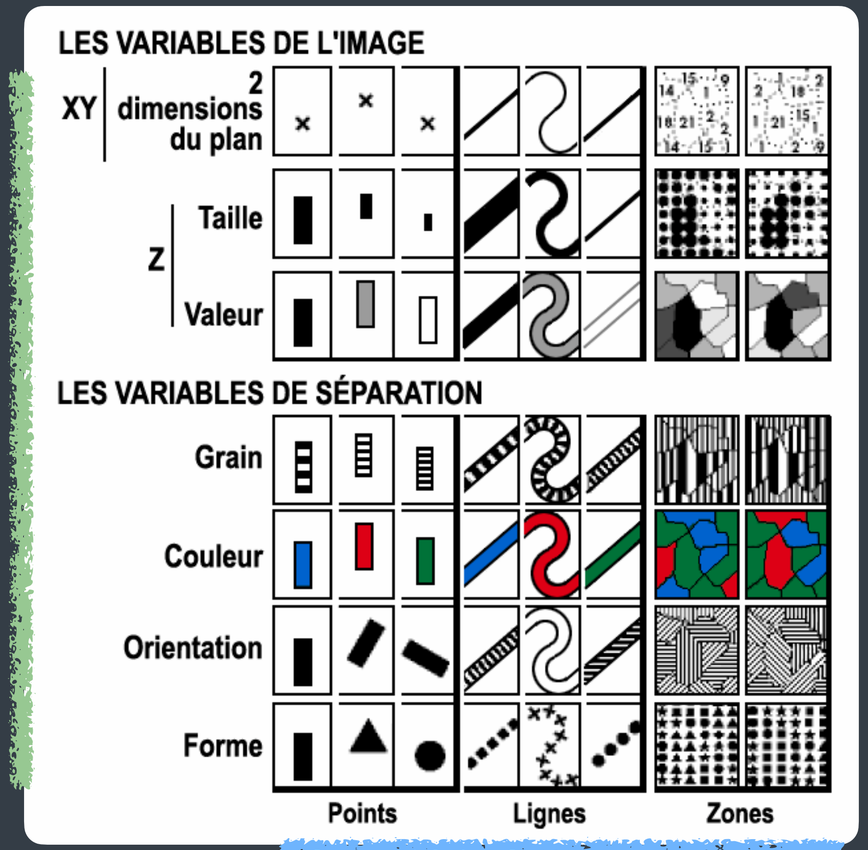
+ friends



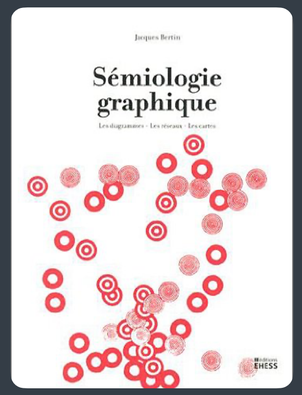


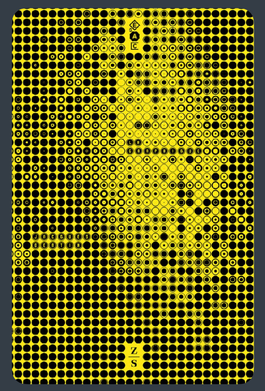
1. A, B, C are distinguishable.
2. B is between A and C.
3. BC is twice as long as AB.

Visual Variables



Marks





Automating the Design of Graphical Presentations of Relational Information

JOCK MACKINLAY
Stanford University

1967

1977

Graphical Perception: Theory, Experimentation, and Application to the Development of Graphical Methods

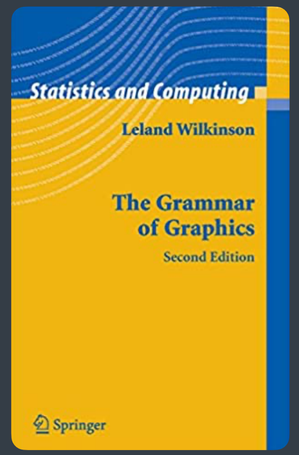
WILLIAM S. CLEVELAND and ROBERT McGILL*

1984

1987

1999

2002

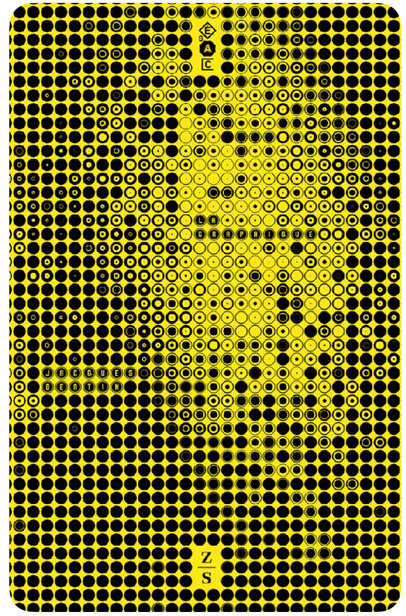


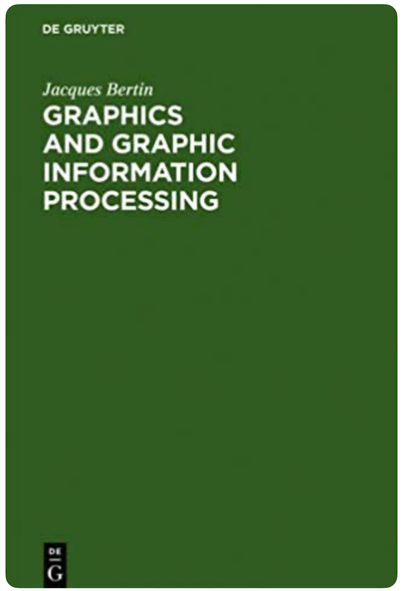
2010

2011

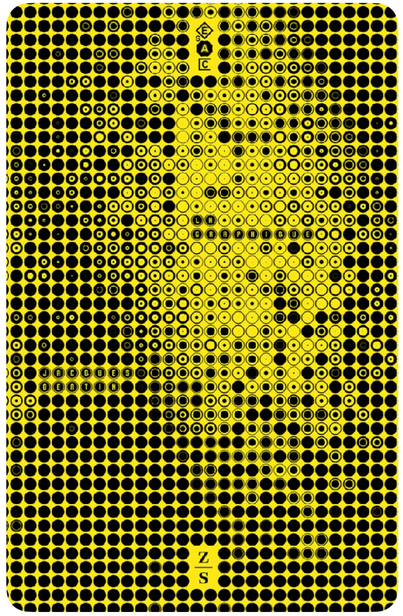
2013

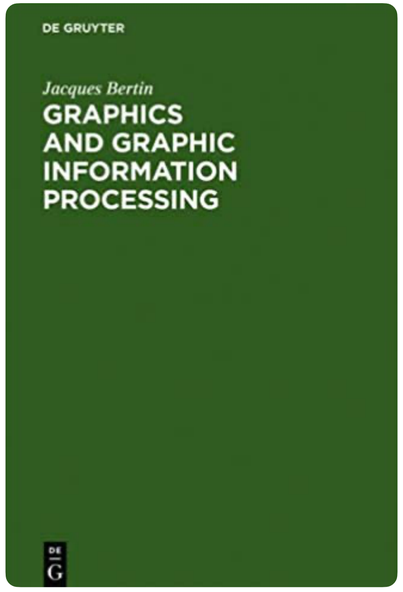
2016

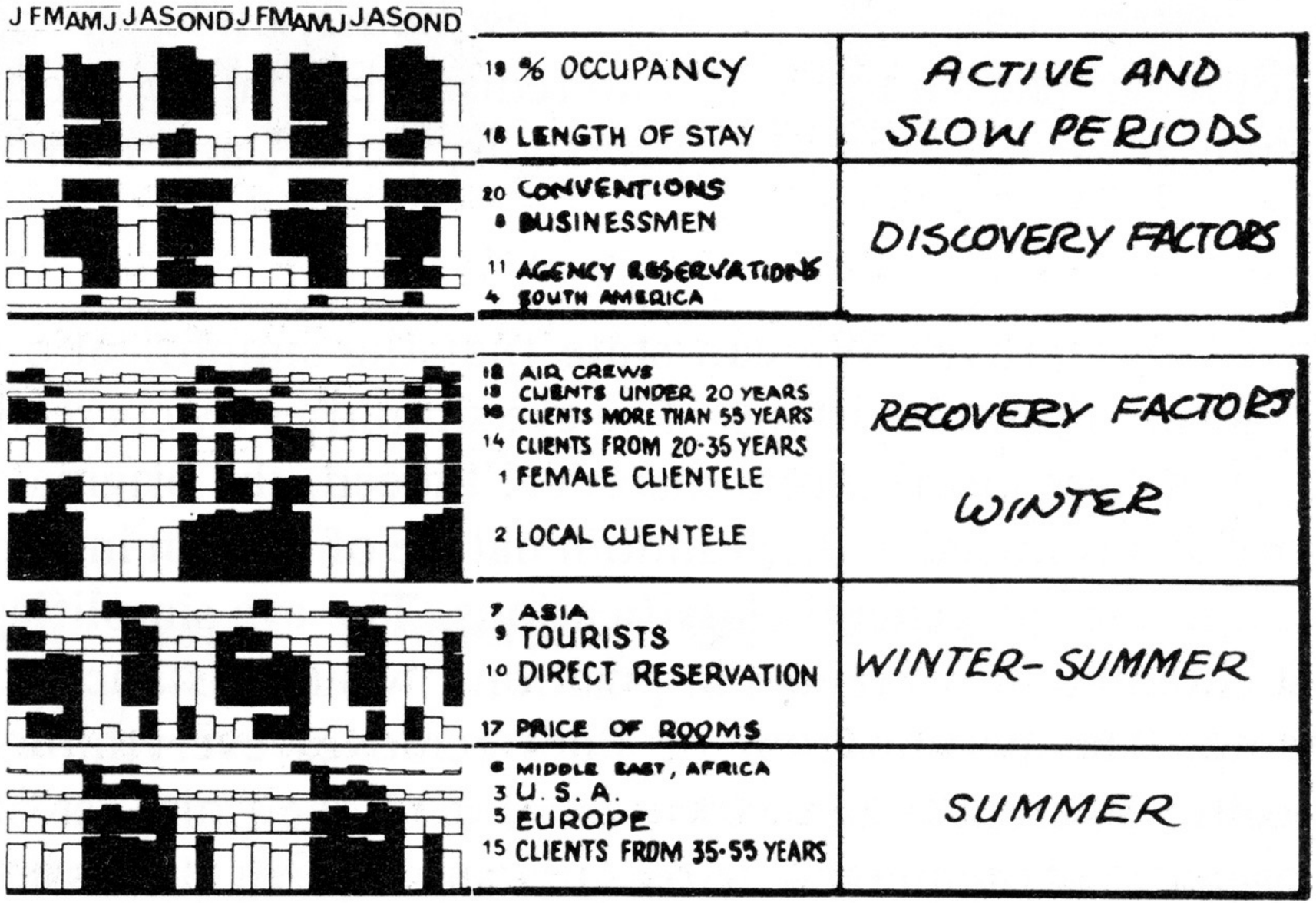


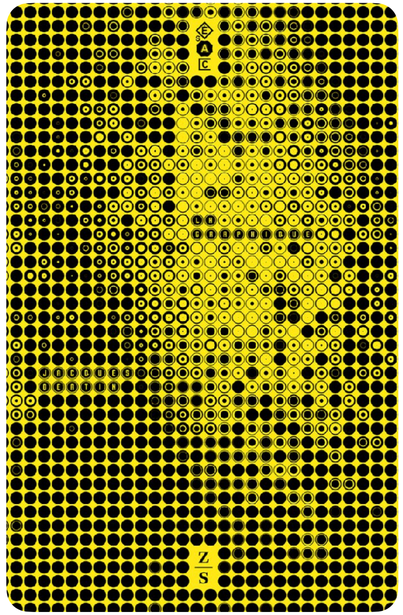


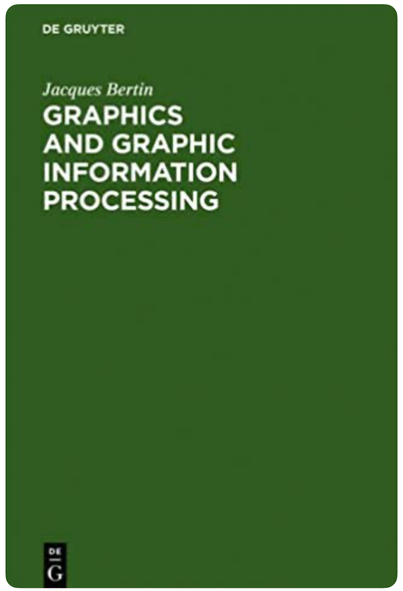
J	F	M	A	M	J	J	A	S	O	N	D		
26	21	26	28	20	20	20	20	20	40	15	40	1	% CLIENTELE FEMALE
69	70	77	71	37	36	39	39	55	60	68	72	2	% —〃— LOCAL
7	6	3	6	23	14	19	14	9	6	8	8	3	% —〃— U.S.A.
0	0	0	0	8	6	6	4	2	12	0	0	4	% —〃— SOUTH AMERICA
20	15	14	15	23	27	22	30	27	19	19	17	5	% —〃— EUROPE
1	0	0	8	6	4	6	4	2	1	0	1	6	% —〃— M.EAST, AFRICA
3	10	6	0	3	13	8	9	5	2	5	2	7	% —〃— ASIA
78	80	85	86	85	87	70	76	87	85	87	80	8	% BUSINESSMEN
22	20	15	14	15	13	30	24	13	15	13	20	9	% TOURISTS
70	70	75	74	69	68	74	75	68	68	64	75	10	% DIRECT RESERVATIONS
20	18	19	17	27	27	19	19	26	27	21	15	11	% AGENCY ———〃———
10	12	6	9	4	5	7	6	6	5	15	10	12	% AIR CREWS
2	2	4	2	2	1	1	2	2	4	2	5	13	% CLIENTS UNDER 20 YEARS
25	27	37	35	25	25	27	28	24	30	24	30	14	% —〃— 20-35 —〃—
48	49	42	48	54	55	53	51	55	46	55	43	15	% —〃— 35-55 —〃—
25	22	17	15	19	19	19	19	19	20	19	22	16	% —〃— MORE THAN 55 —〃—
163	167	166	174	152	155	145	170	157	174	165	156	17	PRICE OF ROOMS
1.65	1.71	1.65	1.91	1.90	2.	1.54	1.60	1.73	1.82	1.66	1.44	18	LENGTH OF STAY
67	82	70	83	74	77	56	62	90	92	78	55	19	% OCCUPANCY
			X	X	X			X	X	X	X	20	CONVENTIONS



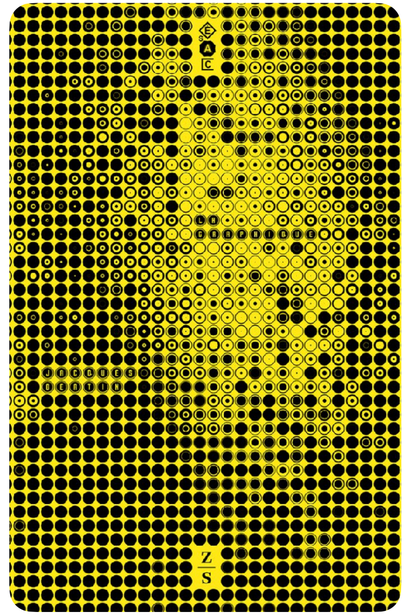

DE GRUYTER
Jacques Bertin
GRAPHICS
AND GRAPHIC
INFORMATION
PROCESSING

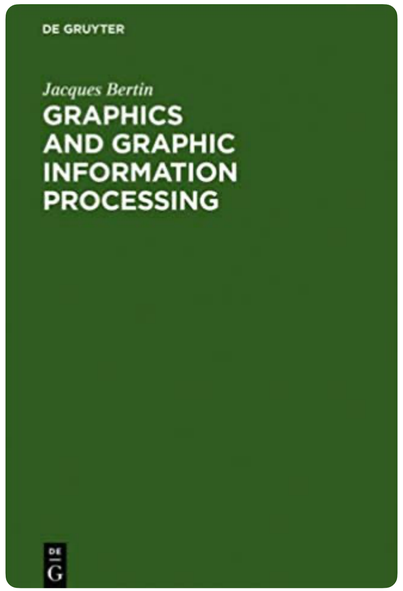

J FMAMJ JASONDJ FMAMJ JASOND
19 % OCCUPANCY
18 LENGTH OF STAY
ACTIVE AND SLOW PERIODS
20 CONVENTIONS
8 BUSINESSMEN
11 AGENCY RESERVATIONS
4 SOUTH AMERICA
DISCOVERY FACTORS
12 AIR CREWS
13 CLIENTS UNDER 20 YEARS
16 CLIENTS MORE THAN 55 YEARS
14 CLIENTS FROM 20-35 YEARS
1 FEMALE CLIENTELE
2 LOCAL CLIENTELE
RECOVERY FACTORS
WINTER
7 ASIA
9 TOURISTS
10 DIRECT RESERVATION
17 PRICE OF ROOMS
WINTER-SUMMER
6 MIDDLE EAST, AFRICA
3 U.S.A.
5 EUROPE
15 CLIENTS FROM 35-55 YEARS
SUMMER

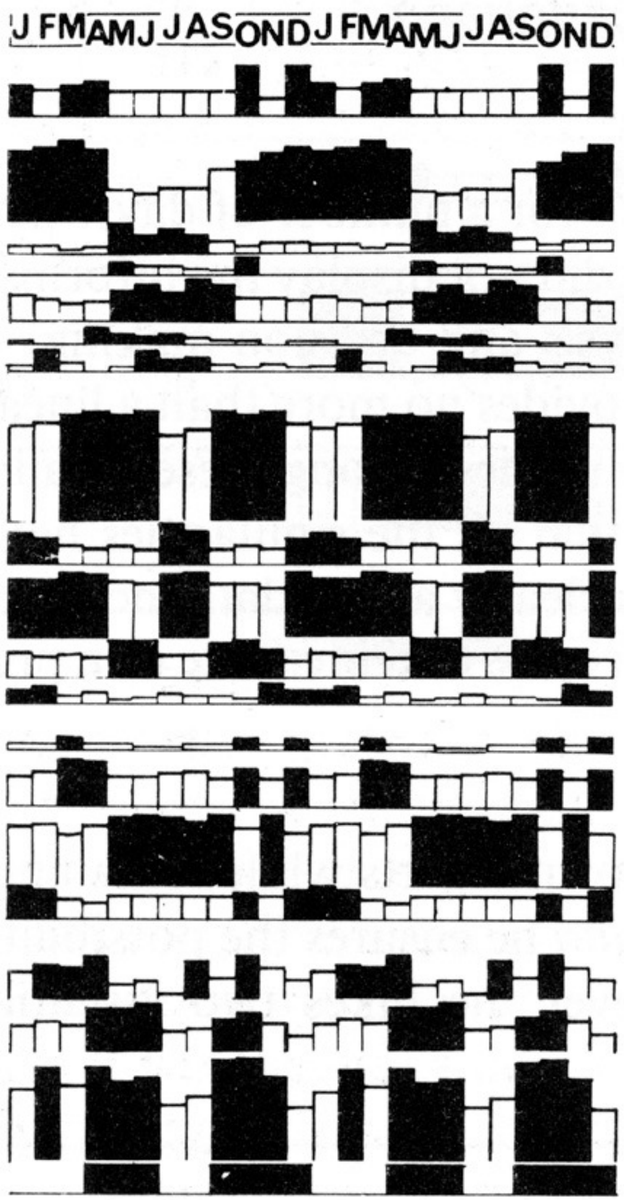




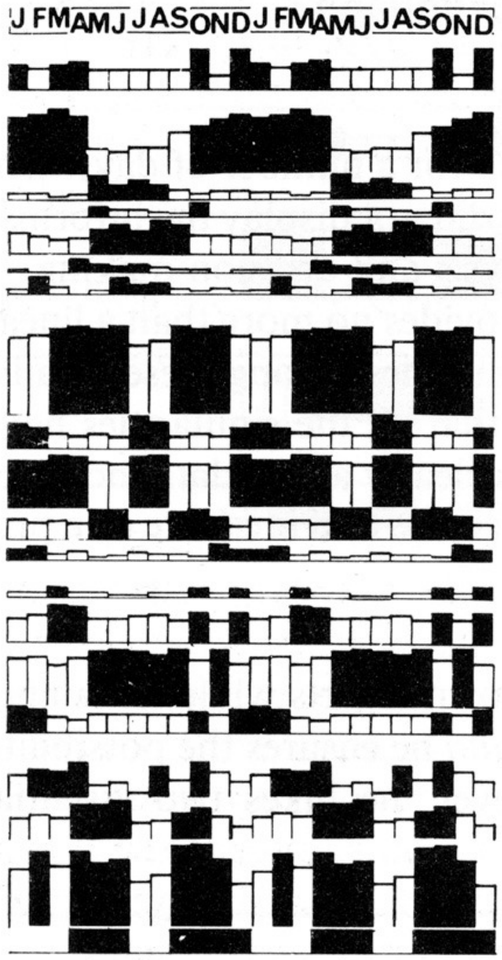
J	F	M	A	M	J	J	A	S	O	N	D		
26	21	26	28	20	20	20	20	20	40	15	40	1	% CLIENTELE FEMALE
69	70	77	71	37	36	39	39	55	60	68	72	2	% —〃— LOCAL
7	6	3	6	23	14	19	14	9	6	8	8	3	% —〃— U.S.A.
0	0	0	0	8	6	6	4	2	12	0	0	4	% —〃— SOUTH AMERICA
20	15	14	15	23	27	22	30	27	19	19	17	5	% —〃— EUROPE
1	0	0	8	6	4	6	4	2	1	0	1	6	% —〃— M.EAST, AFRICA
3	10	6	0	3	13	8	9	5	2	5	2	7	% —〃— ASIA
78	80	85	86	85	87	70	76	87	85	87	80	8	% BUSINESSMEN
22	20	15	14	15	13	30	24	13	15	13	20	9	% TOURISTS
70	70	75	74	69	68	74	75	68	68	64	75	10	% DIRECT RESERVATIONS
20	18	19	17	27	27	19	19	26	27	21	15	11	% AGENCY ——〃——
10	12	6	9	4	5	7	6	6	5	15	10	12	% AIR CREWS
2	2	4	2	2	1	1	2	2	4	2	5	13	% CLIENTS UNDER 20 YEARS
25	27	37	35	25	25	27	28	24	30	24	30	14	% —〃— 20-35 —〃—
48	49	42	48	54	55	53	51	55	46	55	43	15	% —〃— 35-55 —〃—
25	22	17	15	19	19	19	19	19	20	19	22	16	% —〃— MORE THAN 55 —〃—
163	167	166	174	152	155	145	170	157	174	165	156	17	PRICE OF ROOMS
1.65	1.71	1.65	1.91	1.90	2.	1.54	1.60	1.73	1.82	1.66	1.44	18	LENGTH OF STAY
67	82	70	83	74	77	56	62	90	92	78	55	19	% OCCUPANCY
			X	X	X			X	X	X	X	20	CONVENTIONS



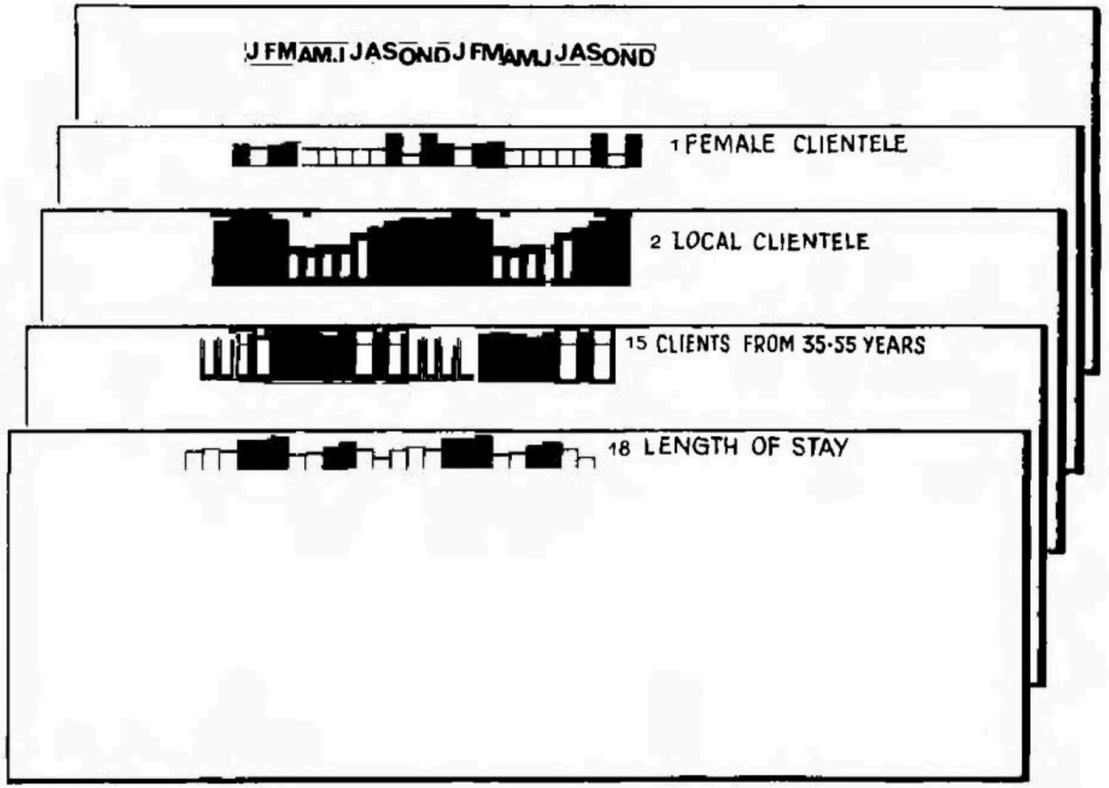


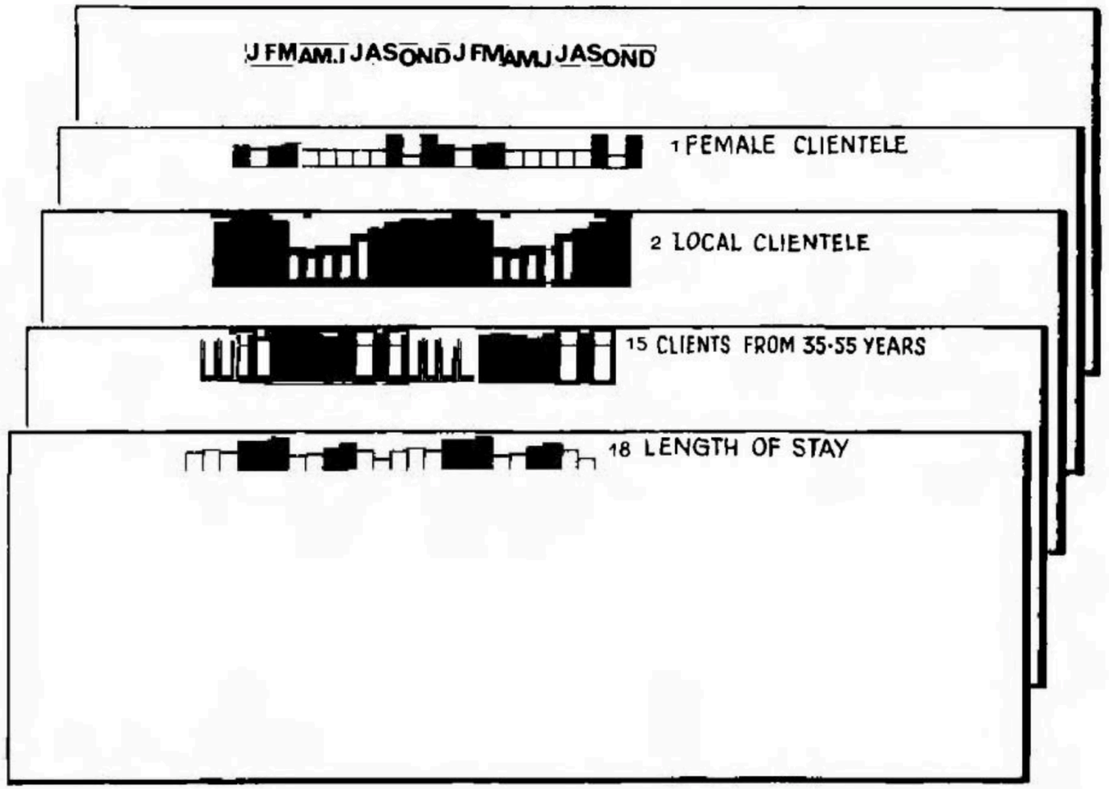


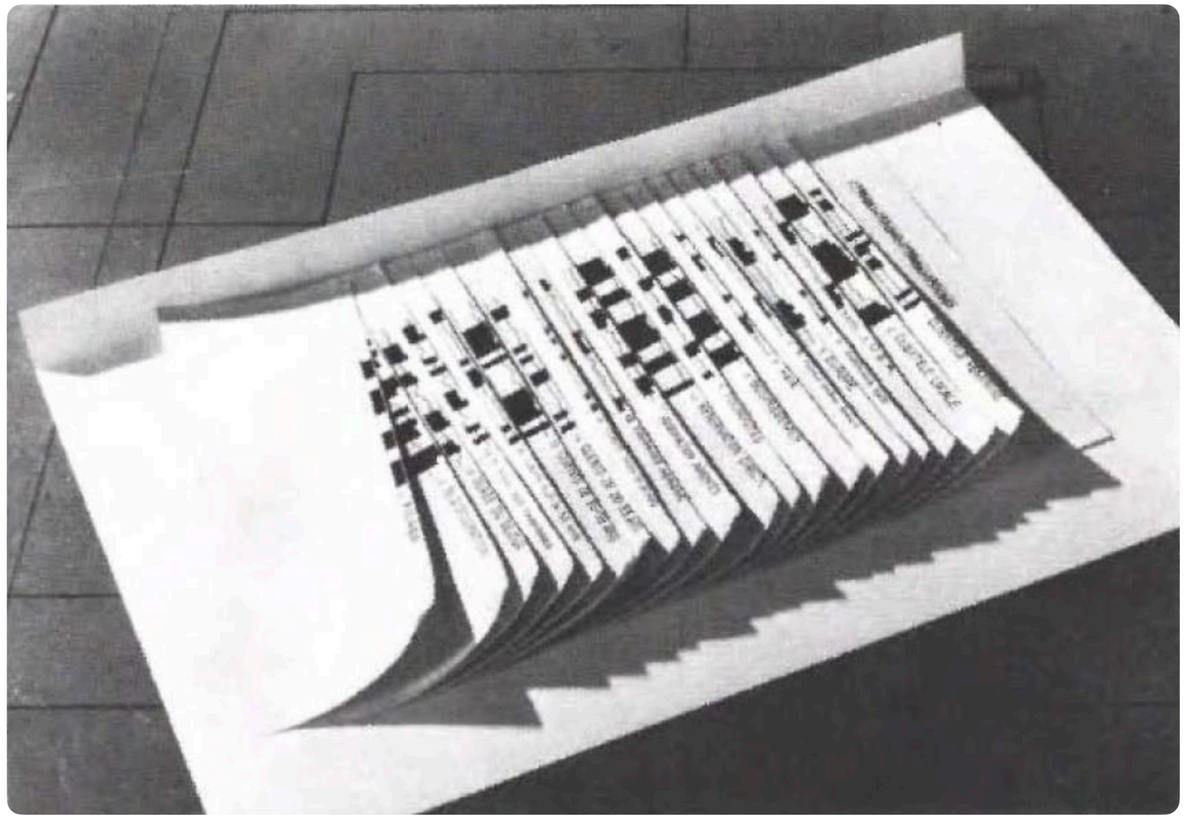
1	% CLIENTELE FEMALE
2	% —〃— LOCAL
3	% —〃— U.S.A.
4	% —〃— SOUTH AMERICA
5	% —〃— EUROPE
6	% —〃— M.EAST, AFRICA
7	% —〃— ASIA
8	% BUSINESSMEN
9	% TOURISTS
10	% DIRECT RESERVATIONS
11	% AGENCY ———〃———
12	% AIR CREWS
13	% CLIENTS UNDER 20 YEARS
14	% —〃— 20-35 —〃—
15	% —〃— 35-55 —〃—
16	% —〃— MORE THAN 55 —〃—
17	PRICE OF ROOMS
18	LENGTH OF STAY
19	% OCCUPANCY
20	CONVENTIONS

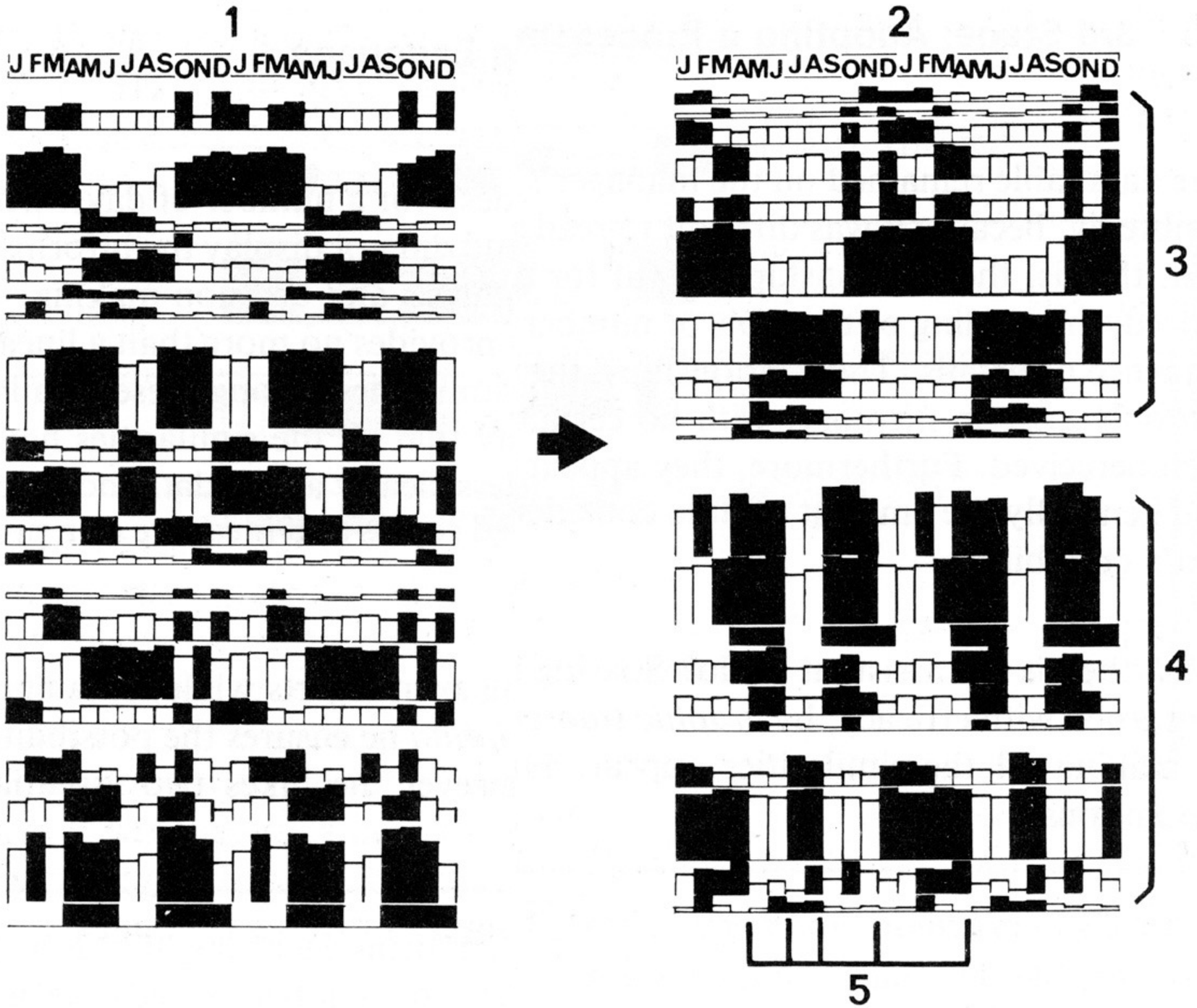


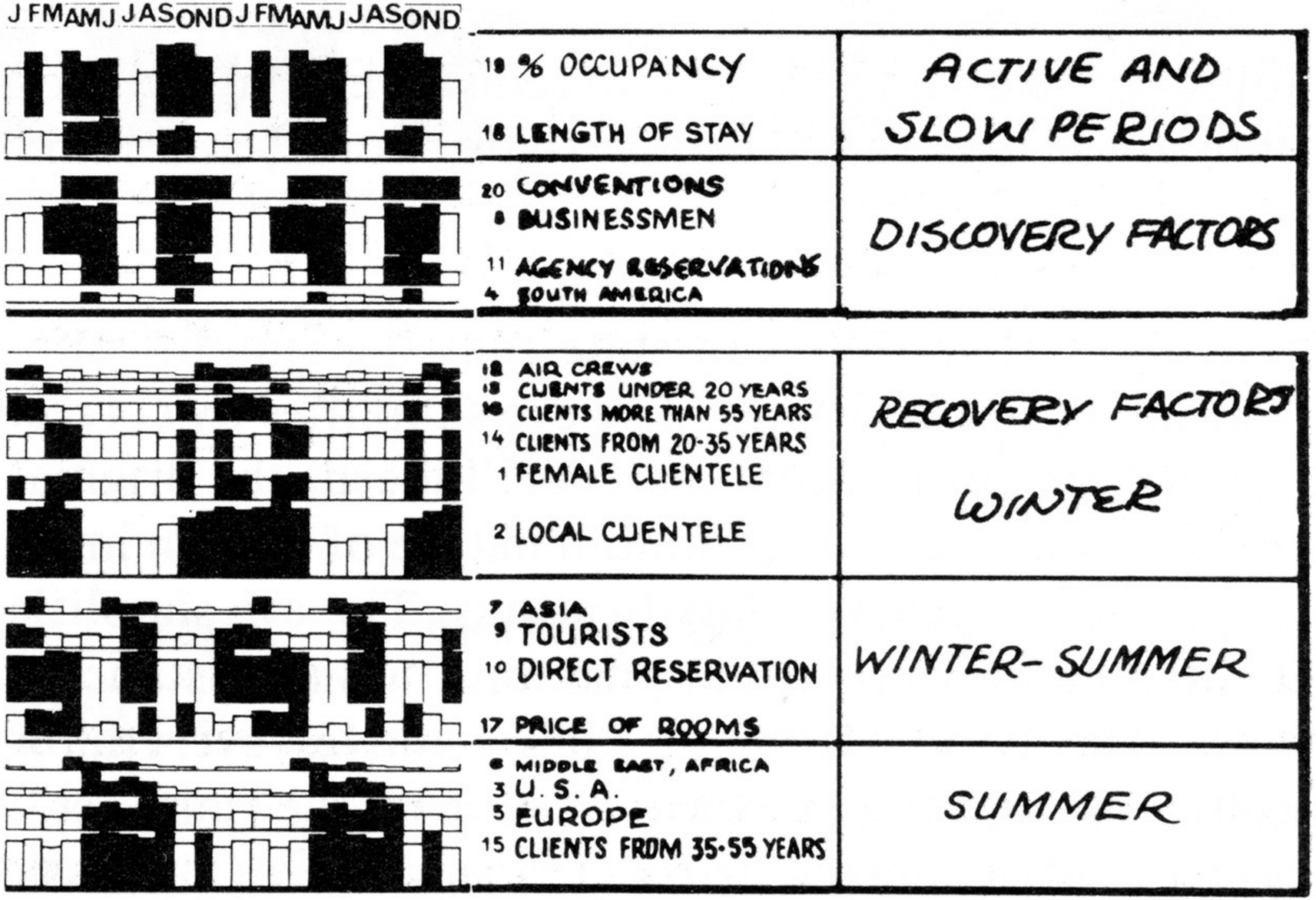
1	% CLIENTELE FEMALE
2	% —//— LOCAL
3	% —//— U.S.A.
4	% —//— SOUTH AMERICA
5	% —//— EUROPE
6	% —//— M.EAST, AFRICA
7	% —//— ASIA
8	% BUSINESSMEN
9	% TOURISTS
10	% DIRECT RESERVATIONS
11	% AGENCY ———//———
12	% AIR CREWS
13	% CLIENTS UNDER 20 YEARS
14	% —//— 20-35 —//—
15	% —//— 35-55 —//—
16	% —//— MORE THAN 55 —//—
17	PRICE OF ROOMS
18	LENGTH OF STAY
19	% OCCUPANCY
20	CONVENTIONS

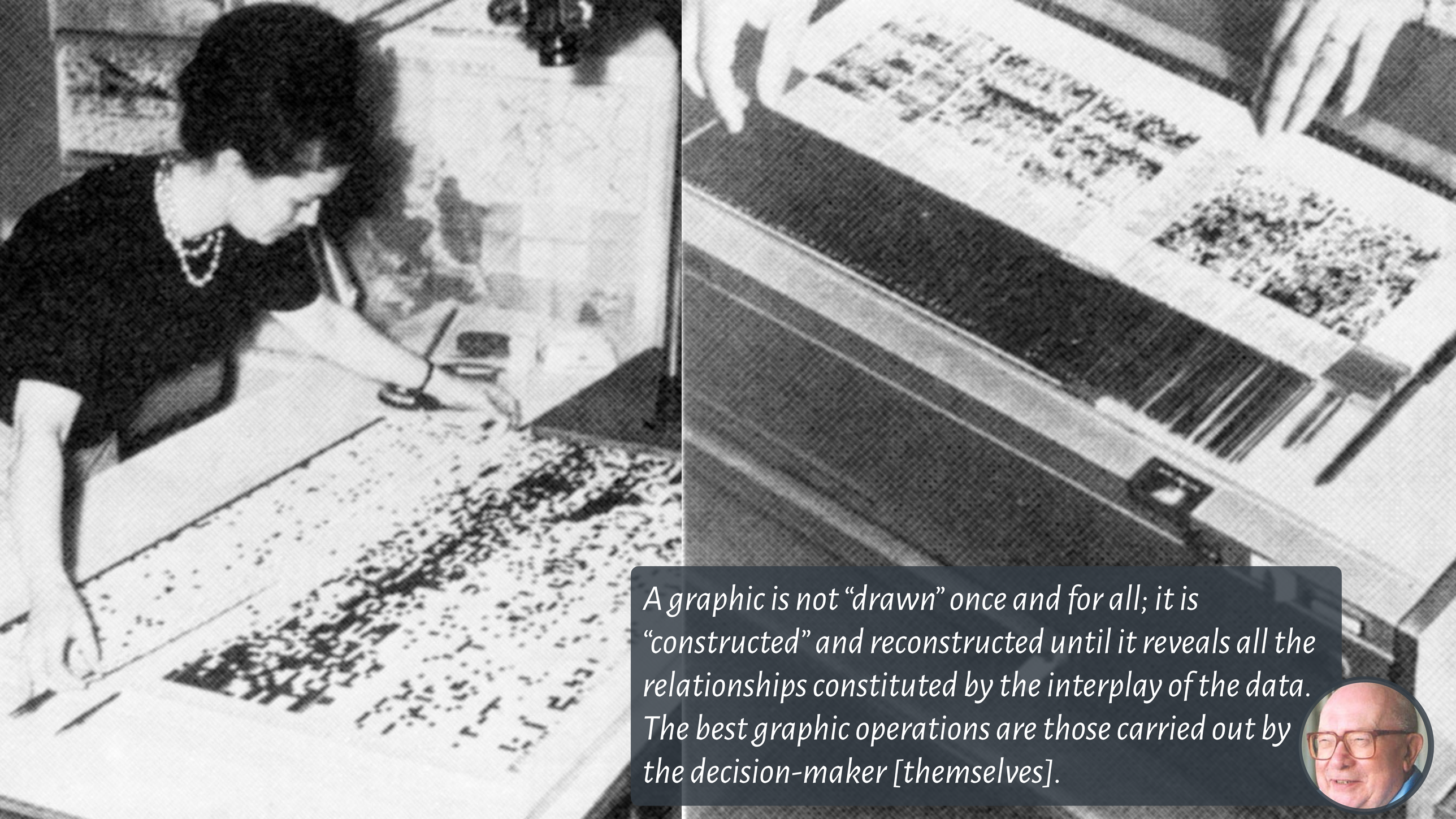



J F M A M J J A S O N D J F M A M J J A S O N D
1 FEMALE CLIENTELE
2 LOCAL CLIENTELE
15 CLIENTS FROM 35-55 YEARS
18 LENGTH OF STAY




1
J FMAMJ JASOND J FMAMJ JASOND
2
J FMAMJ JASOND J FMAMJ JASOND
3
4
5

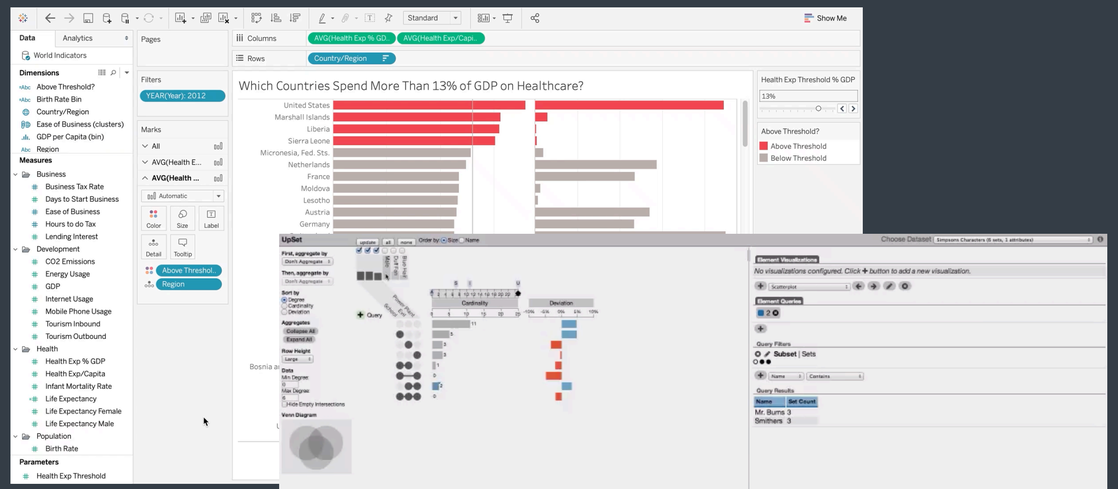

J FMAMJ JASOND J FMAMJ JASOND
19 % OCCUPANCY
18 LENGTH OF STAY
ACTIVE AND SLOW PERIODS
20 CONVENTIONS
8 BUSINESSMEN
11 AGENCY RESERVATIONS
4 SOUTH AMERICA
DISCOVERY FACTORS
12 AIR CREWS
13 CLIENTS UNDER 20 YEARS
16 CLIENTS MORE THAN 55 YEARS
14 CLIENTS FROM 20-35 YEARS
1 FEMALE CLIENTELE
2 LOCAL CLIENTELE
RECOVERY FACTORS WINTER
7 ASIA
9 TOURISTS
10 DIRECT RESERVATION
17 PRICE OF ROOMS
WINTER-SUMMER
6 MIDDLE EAST, AFRICA
3 U.S.A.
5 EUROPE
15 CLIENTS FROM 35-55 YEARS
SUMMER

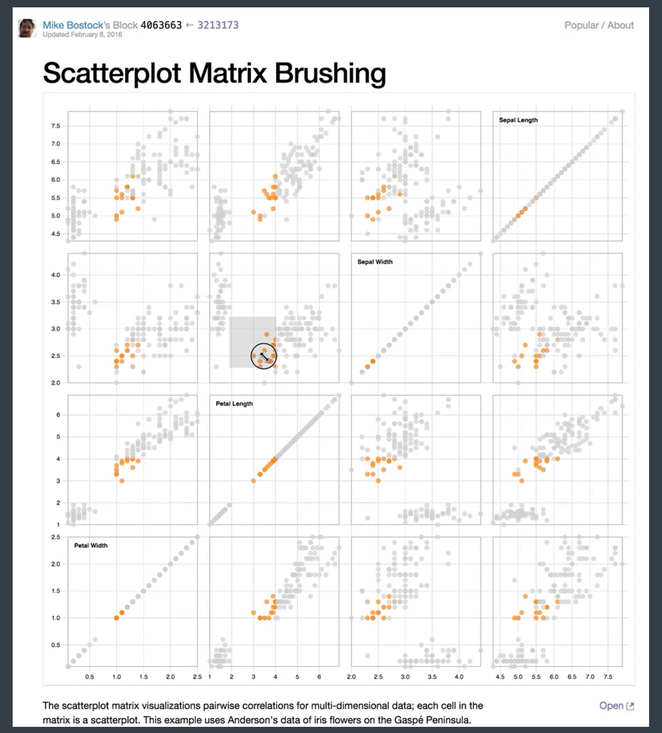


A graphic is not "drawn" once and for all; it is "constructed" and reconstructed until it reveals all the relationships constituted by the interplay of the data. The best graphic operations are those carried out by the decision-maker [themselves].

"Visual representations alone cannot satisfy analytical needs. Interaction techniques are required to support the dialogue between the analyst and the data."
[Thomas & Cook, 2005]

"It is through the interactive manipulation of a visual interface—the analytic discourse—that knowledge is constructed, tested, refined, and shared."
[Pike et al., 2009]





???

Fixed palette of general-purpose techniques

Custom techniques in purpose-built tools

Programming custom techniques

Vega-Lite: A Grammar of Interactive Graphics

with   

```
{
  "data": {"url": "data/gapminder.json"},
  "mark": "circle",
  "encoding": {
    "x": {"field": "fertility","type": "Q"},
    "y": {"field": "life_expectancy","type": "Q"},
    "color": {"field": "region","type": "N"},
    "size": {"field": "population","type": "Q"},
  }
}
```

Vega-Lite: A Grammar of Interactive Graphics

with   

```
{
  "data": {"url": "data/gapminder.json"},
  "mark": "circle",
  "encoding": {
    "x": {"field": "fertility","type": "Q"},
    "y": {"field": "life_expectancy","type": "Q"},
    "color": {"field": "region","type": "N"},
    "size": {"field": "population","type": "Q"},
  }
}
```

Altair (Python)

```
alt.Chart("data/gapminder.json")
  .mark_circle()
  .encode(
    x='fertility:Q',
    y='life_expectancy:Q',
    color='region:N',
    size='population:Q'
  )
```

vegalite (R)

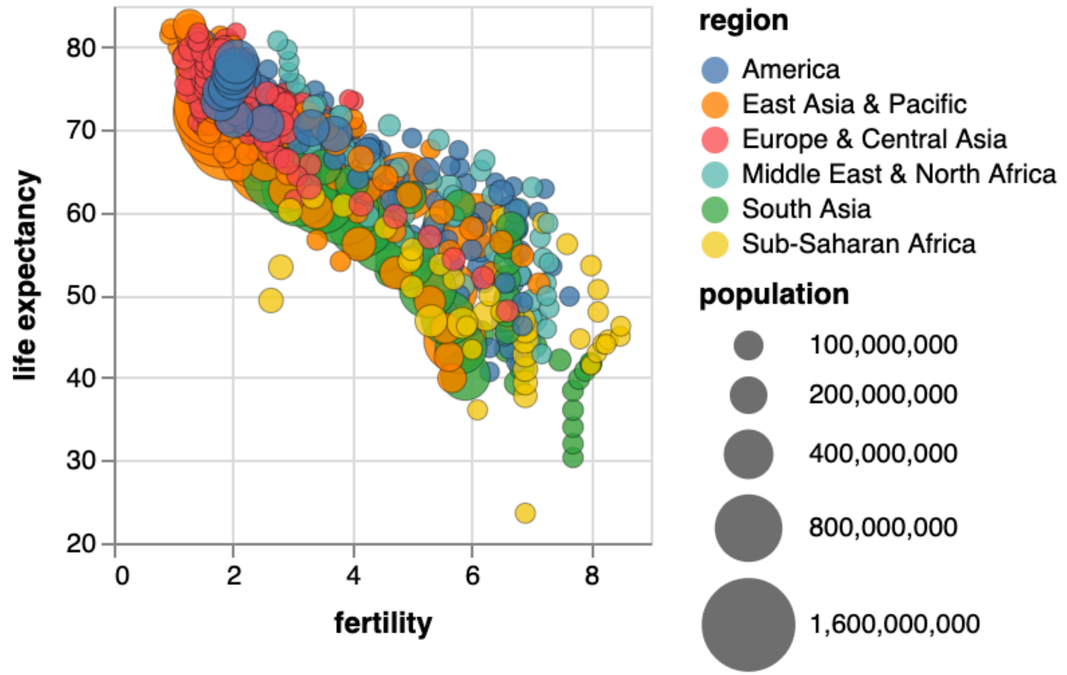
```
vegalite() %>%
  add_data("data/gapminder.json") %>%
  encode_x("fertility", "Q") %>%
  encode_y("life_expectancy", "Q") %>%
  encode_color("region", "N") %>%
  encode_size("population", "Q") %>%
  mark_circle()
```

Vega-Lite: A Grammar of Interactive Graphics

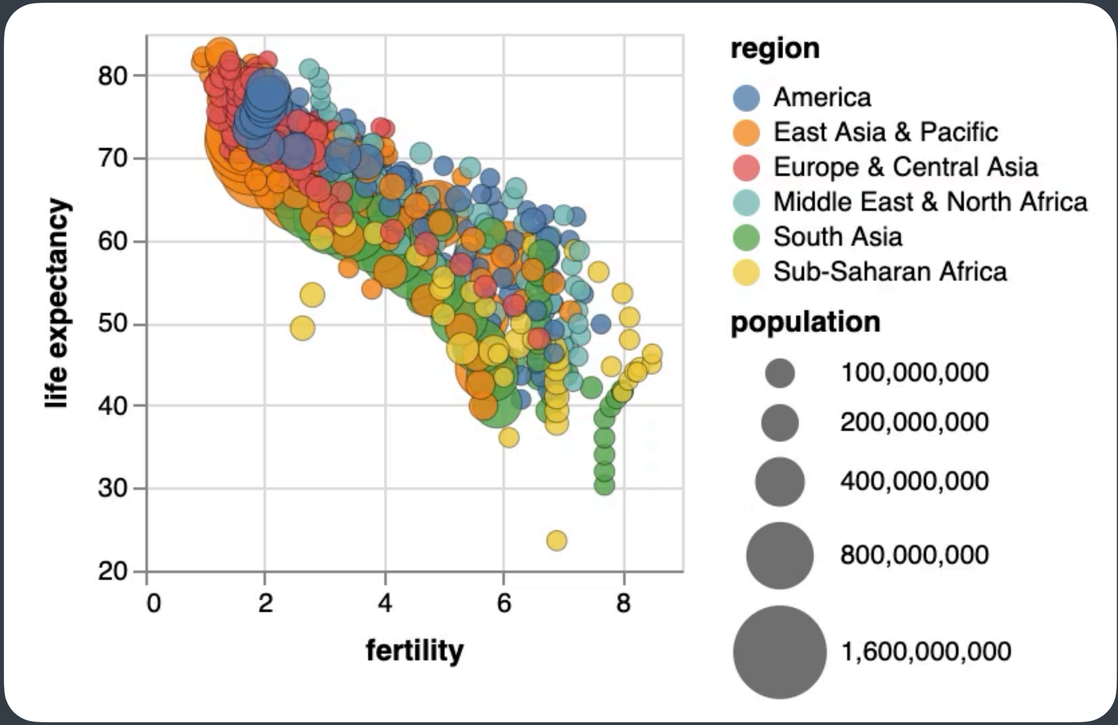
with



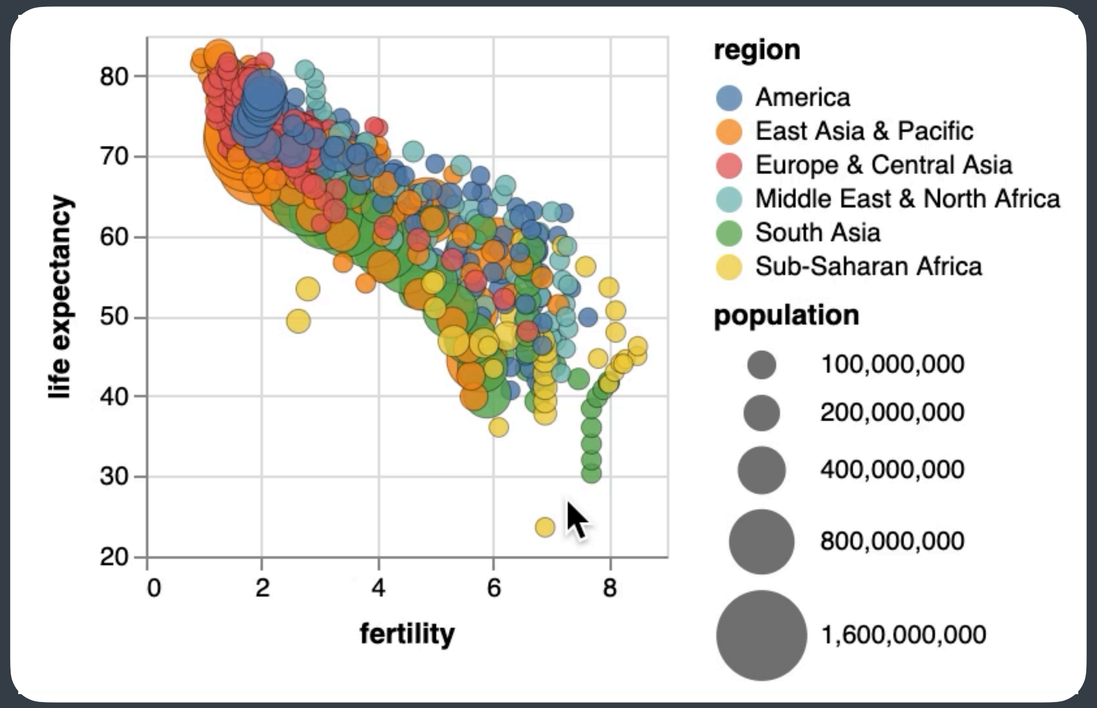
```
{
  "data": {"url": "data/gapminder.json"},
  "mark": "circle",
  "encoding": {
    "x": {"field": "fertility","type": "Q"},
    "y": {"field": "life_expectancy","type": "Q"},
    "color": {"field": "region","type": "N"},
    "size": {"field": "population","type": "Q"},
  }
}
```



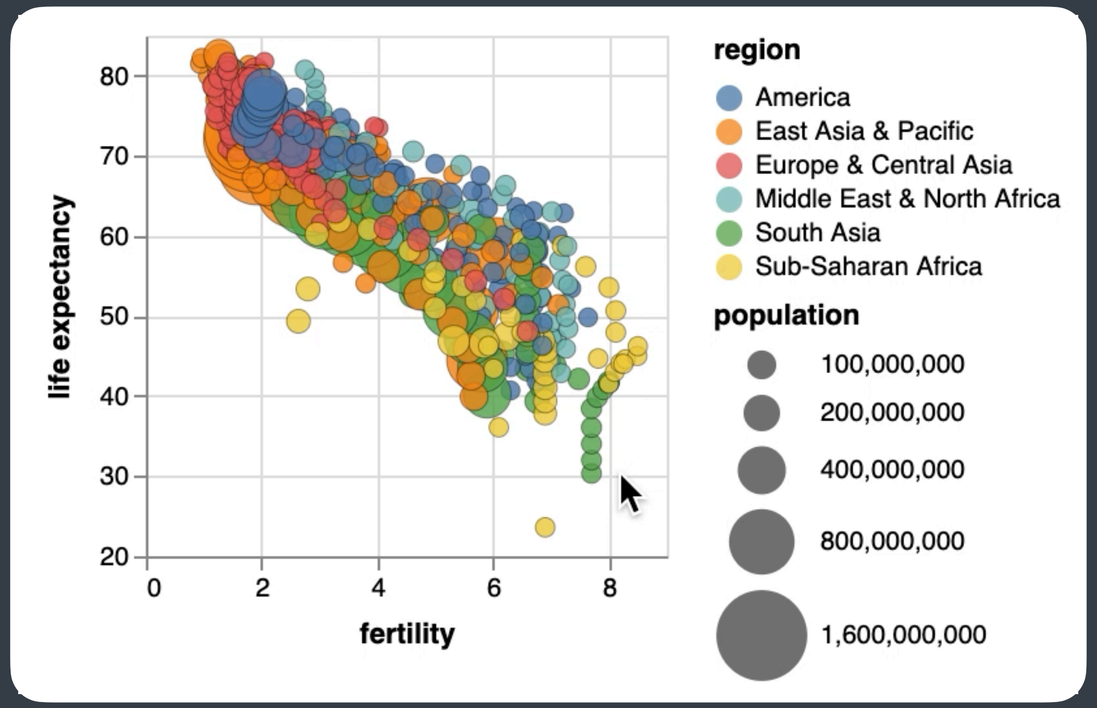
```
{
  "data": {"url": "data/gapminder.json"},
  "mark": "circle",
  "encoding": {
    "x": {"field": "fertility","type": "Q"},
    "y": {"field": "life_expectancy","type": "Q"},
    "color": {"field": "region","type": "N"},
    "size": {"field": "population","type": "Q"},
    "tooltip": [
      {"field": "country"},
      {"field": "year"}
    ]
  }
}
```



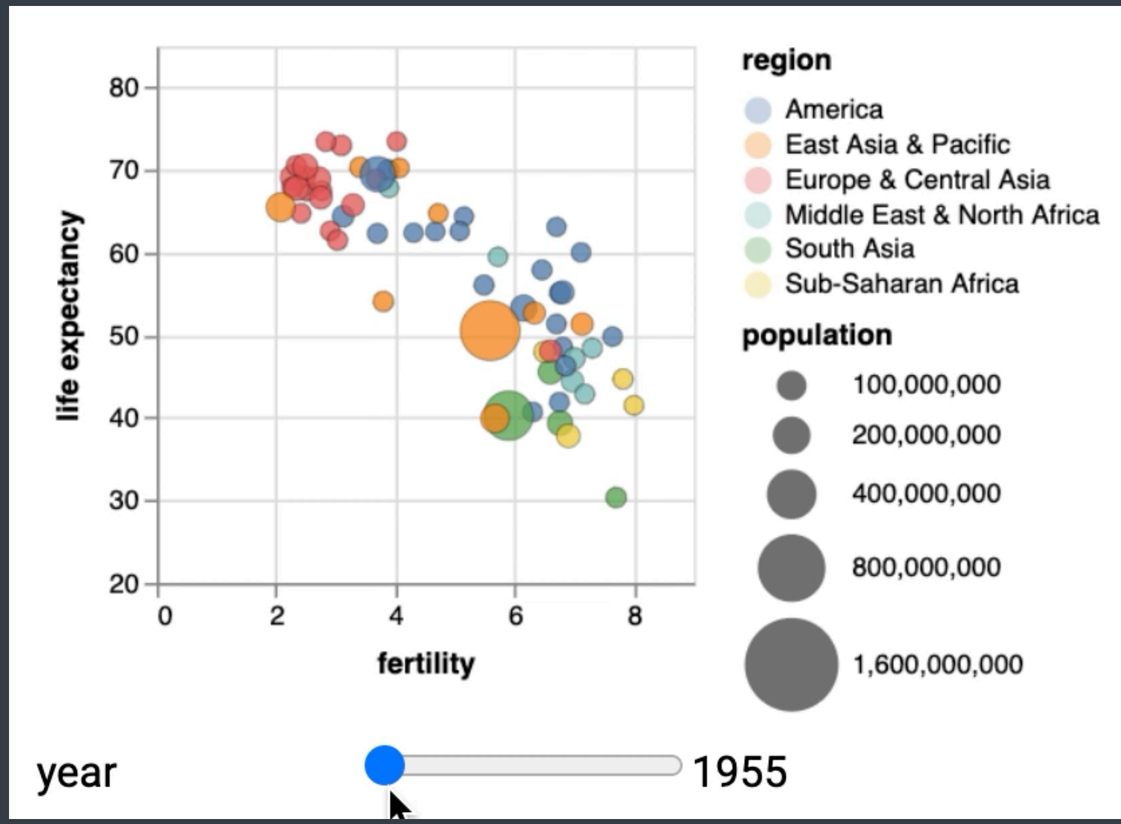
```
{
  "data": {"url": "data/gapminder.json"},
  "mark": "circle",
  "selection": {
    "cnty": {"type": "single","fields": ["country"]}
  },
  "encoding": {
    "x": {"field": "fertility","type": "Q"},
    "y": {"field": "life_expectancy","type": "Q"},
    "color": {"field": "region","type": "N"},
    "size": {"field": "population","type": "Q"},
    "opacity": {
      "condition": {"selection": "cnty","value": 1},
      "value": 0.25
    },
    "tooltip": [
      {"field": "country"},
      {"field": "year"}
    ],
  }
}
```



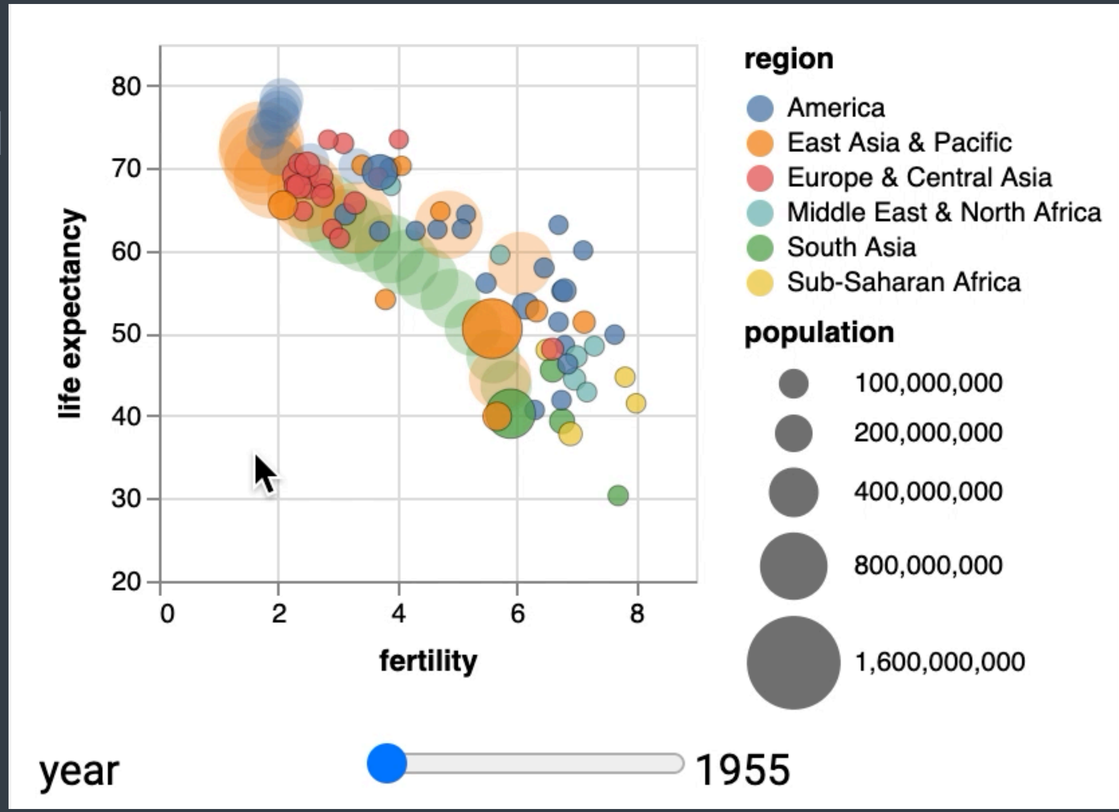
```
{
  "data": {"url": "data/gapminder.json"},
  "mark": "circle",
  "selection": {
    "cnty": {"type": "multi" ,"fields": ["country"]}
  },
  "encoding": {
    "x": {"field": "fertility","type": "Q"},
    "y": {"field": "life_expectancy","type": "Q"},
    "color": {"field": "region","type": "N"},
    "size": {"field": "population","type": "Q"},
    "opacity": {
      "condition": {"selection": "cnty","value": 1},
      "value": 0.25
    },
    "tooltip": [
      {"field": "country"},
      {"field": "year"}
    ],
  }
}
```



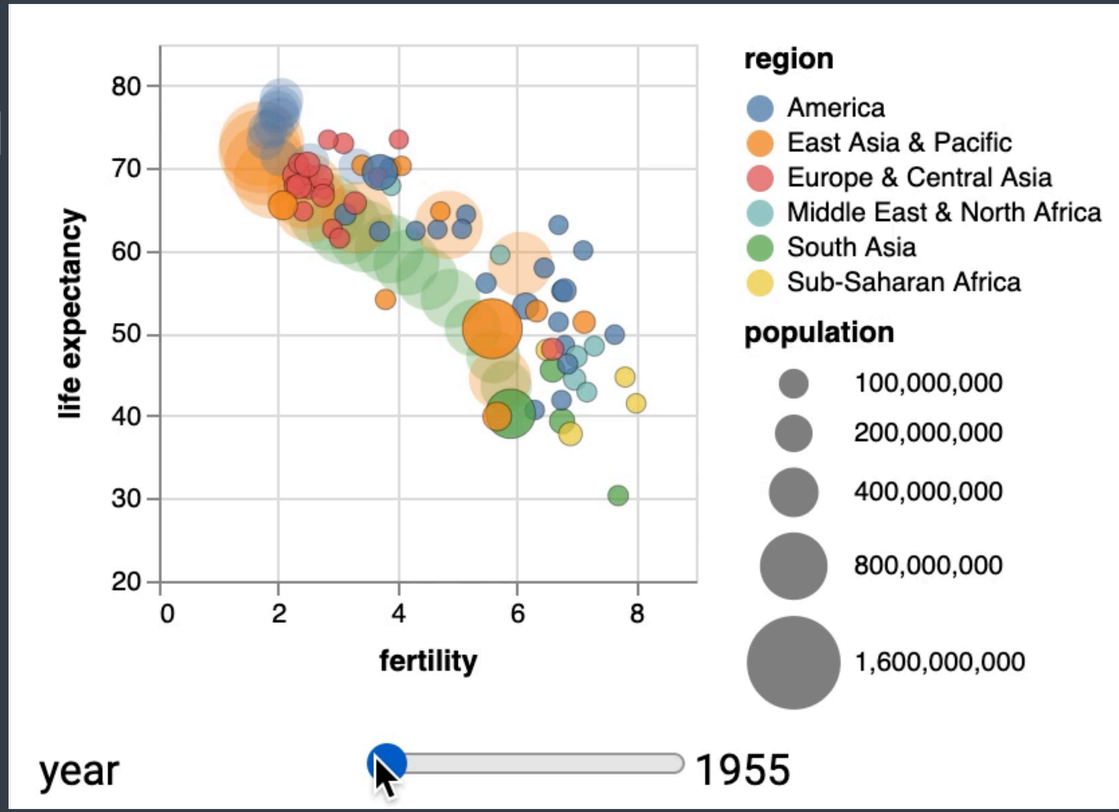
```
{
  "data": {"url": "data/gapminder.json"},
  "transform": [{"filter": {"selection": "yr"}}],
  "mark": "circle",
  "selection": {
    "cnty": {"type": "multi" ,"fields": ["country"]}
    "yr": {
      "type": "single","fields": ["year"]
      "bind": {"input": "range", ...}
    }
  },
  "encoding": {
    "x": {"field": "fertility","type": "Q"},
    "y": {"field": "life_expectancy","type": "Q"},
    "color": {"field": "region","type": "N"},
    "size": {"field": "population","type": "Q"},
    "tooltip": [
      {"field": "country"},
      {"field": "year"}
    ],
  }
}
```



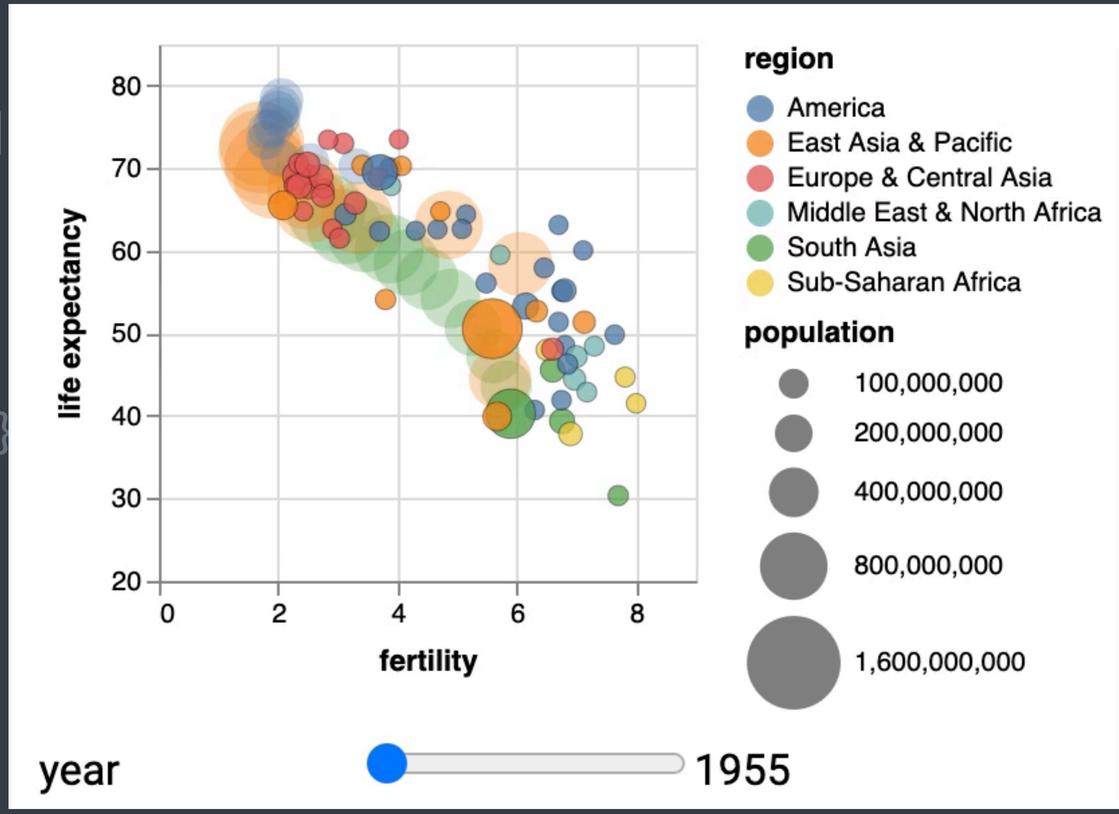
```
{
  "data": {"url": "data/gapminder.json"},
  "layer": [{
    "transform": [{"filter": {"selection": "yr"}}],
    "mark": "circle",
    "selection": {
      "cnty": {"type": "multi"  ,"fields": ["country"]
      "yr": {
        "type": "single","fields": ["year"]
        "bind": {"input": "range", ...}
      }
    },
    "encoding": {...}
  }, {
    "mark": "circle",
    "encoding": {
      ...,
      "opacity": {
        "condition": {"selection": "cnty","value": 1},
        "value": 0.25
      },
    }
  }]
```



```
{
  "data": {"url": "data/gapminder.json"},
  "layer": [{
    "transform": [{"filter": {"selection": "yr"}}],
    "mark": "circle",
    "selection": {
      "cnty": {"type": "multi"  ,"fields": ["country"]
      "yr": {
        "type": "single","fields": ["year"]
        "bind": {"input": "range", ...}
      },
      "brush": {"type": "interval"}
  },
    "encoding": {...}
  }, {
    "mark": "circle",
    "encoding": {
      ...,
      "color": {
        "condition": {"selection": "brush","field": "region","type": "N"},
        "value": "grey"
      },
    }
```

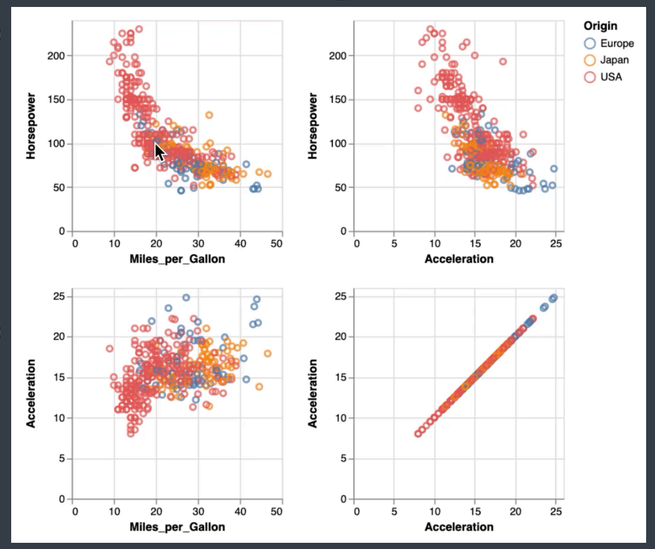


```
{
  "data": {"url": "data/gapminder.json"},
  "layer": [{
    "transform": [{"filter": {"selection": "yr"}}],
    "mark": "circle",
    "selection": {
      "cnty": {"type": "multi"  ,"fields": ["country"]
      "yr": {
        "type": "single","fields": ["year"]
        "bind": {"input": "range", ...}
      },
      "brush": {"type": "interval", "bind": "scales"}
  },
    "encoding": {...}
  }, {
    "mark": "circle",
    "encoding": {
      ...,
      "color": {
        "condition": {"selection": "brush","field": "region","type": "N"},
        "value": "grey"
      },
    }
```

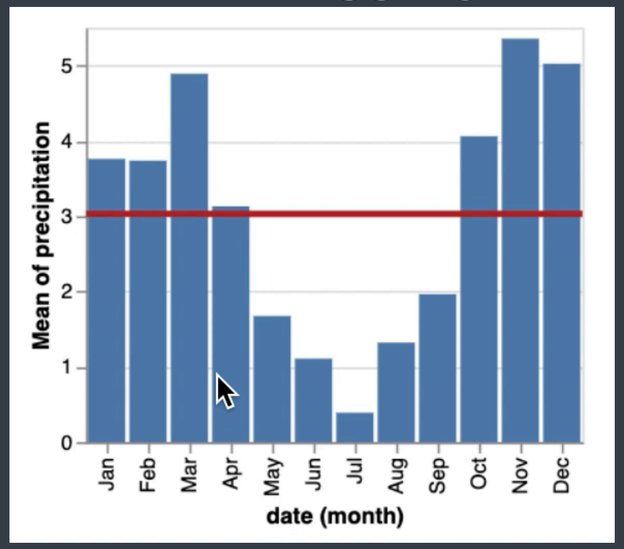


Panning & Zooming

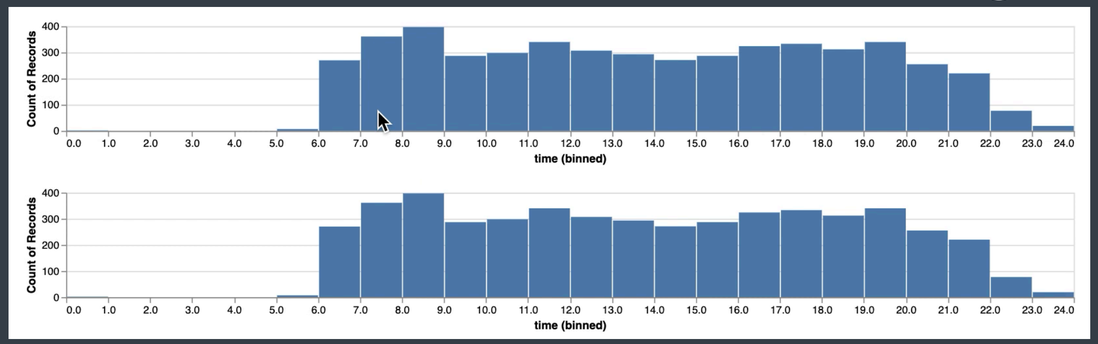
Union Multiple Brushes



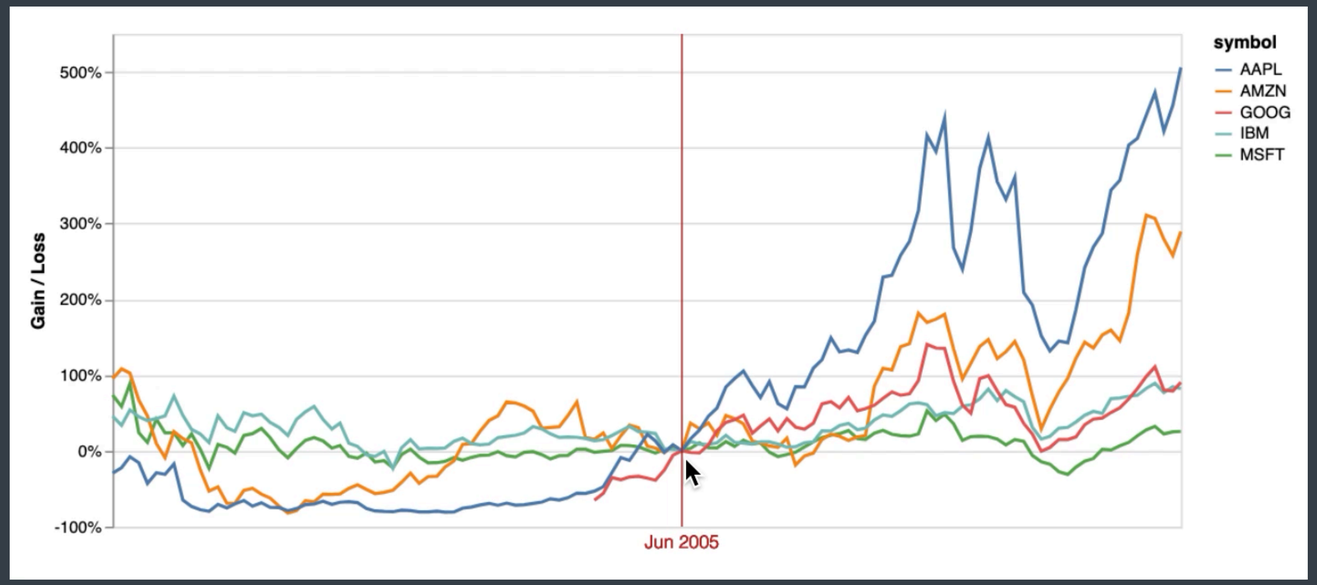
Interactive Aggregation



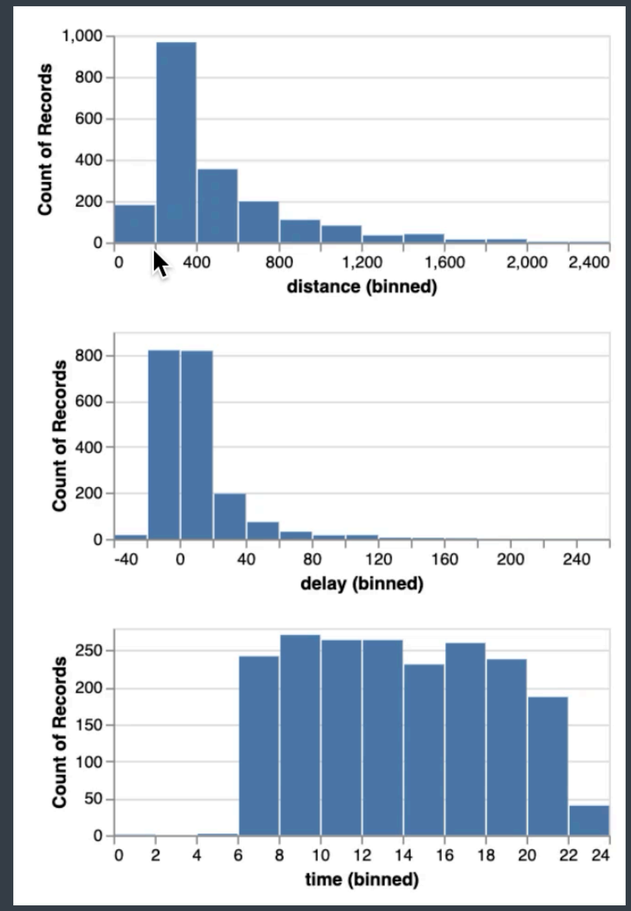
Overview+Detail Interactive Binning



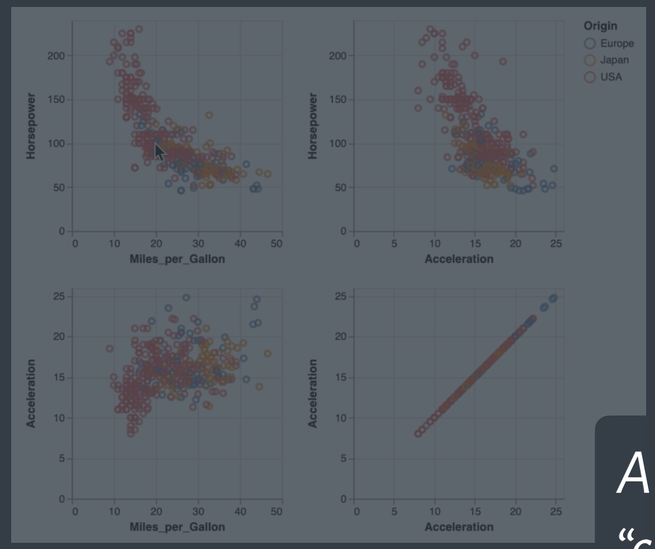
Interactive Re-Normalization



Interactive Cross Filtering

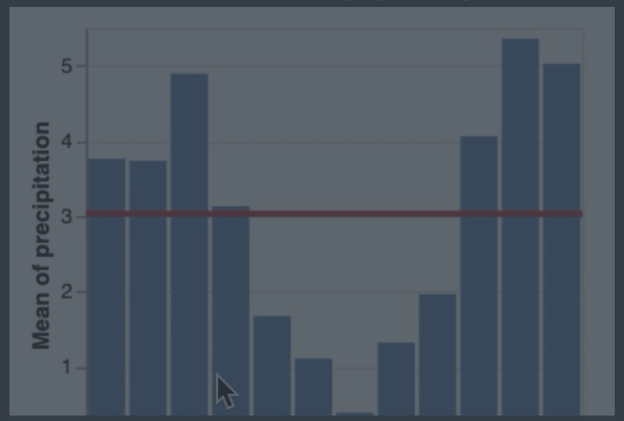


Union Multiple Brushes

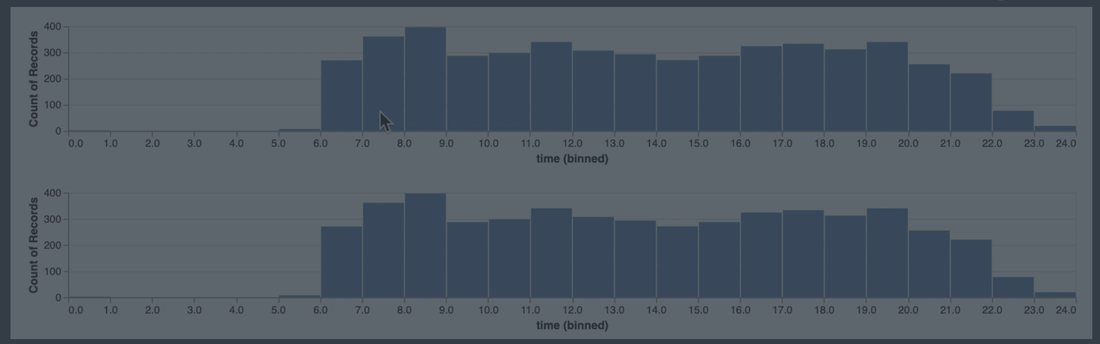


Panning & Zooming

Interactive Aggregation

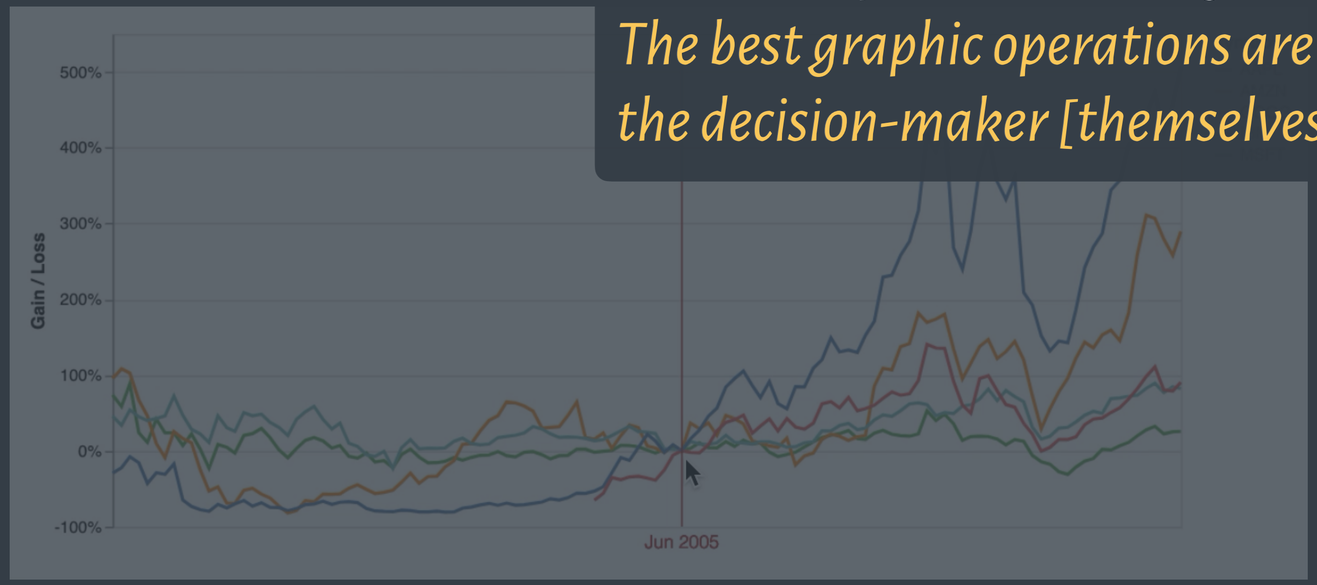


Overview+Detail Interactive Binning

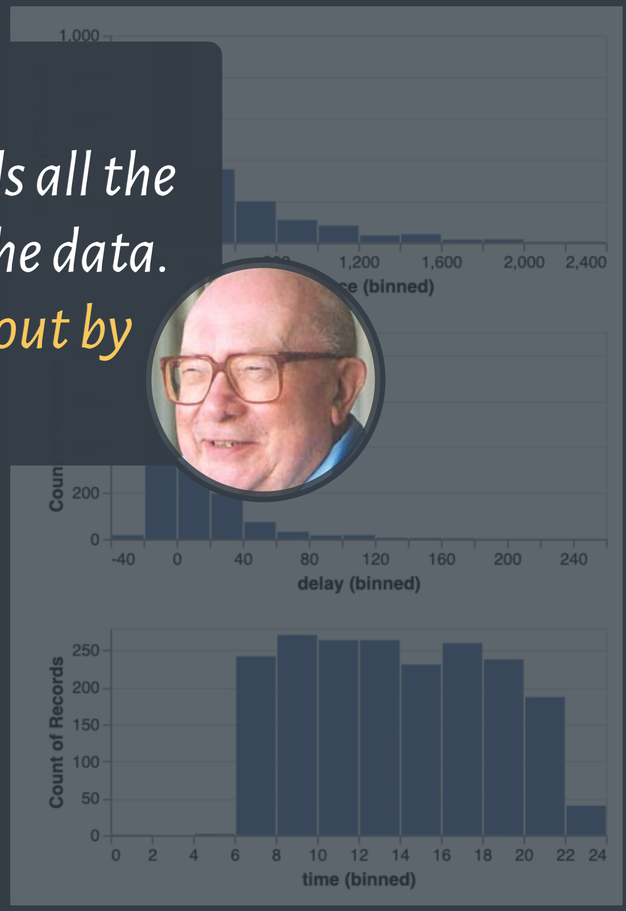


A graphic is not "drawn" once and for all; it is "constructed" and reconstructed until it reveals all the relationships constituted by the interplay of the data.
The best graphic operations are those carried out by the decision-maker [themselves].

Interactive F



Interactive Cr



A *Layout* Gap

Visualizations interleaved with code makes it difficult to coordinate multiple views.

A *Semantic* Gap

Visualizations need to be manually specified despite the provenance expressed by dataframe transformations.

Interactive results are siloed, and not available for further analysis in code.

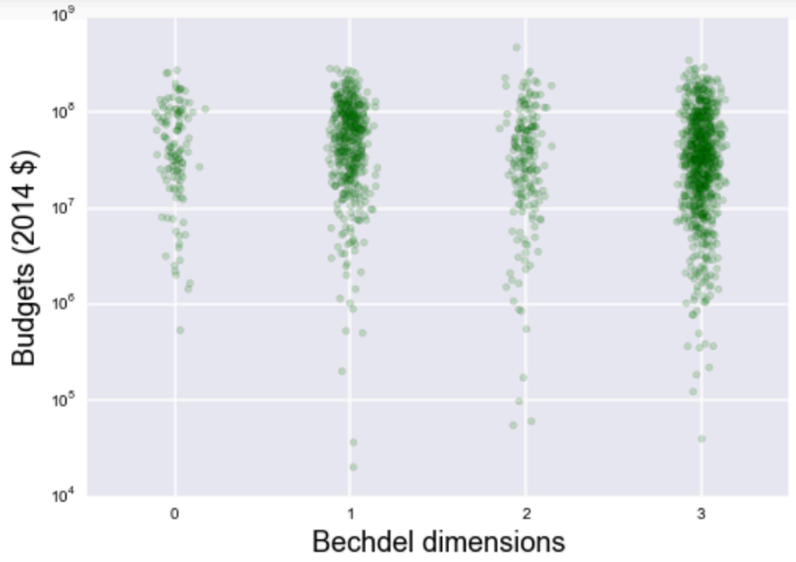
A *Temporal* Gap

Cell execution is persistent, but interaction is transient.

jupyter Bechdel_test (autosaved) Logout

File Edit View Insert Cell Kernel Help Not Trusted Kernel

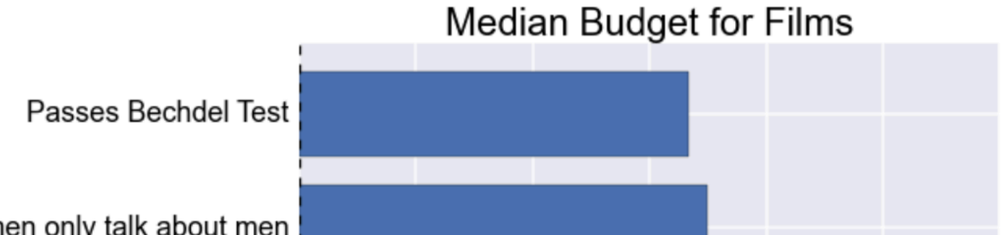
Run Markdown



We can also visualize the median budgets with a bar chart to see that movies featuring two women who don't talk to each other appear to have much larger budgets than the rest. Movies that pass the Bechdel test also appear to have slightly smaller budgets than movies that don't pass.

In [43]:
```
df['Adj_Budget'].groupby(df['rating']).agg(np.median).plot(kind='barh')
plt.xticks(arange(0e7,7e7,1e7),arange(0,70,10),fontsize=15)
plt.yticks(plt.yticks()[0],["Fewer than two women",
                            "Women don't talk to each other",
                            'Women only talk about men',
                            'Passes Bechdel Test'
                            ],fontsize=18)
plt.xlabel('Budget (2014 $millions)',fontsize=18)
plt.title('Median Budget for Films',fontsize=24)
plt.ylabel('')
```

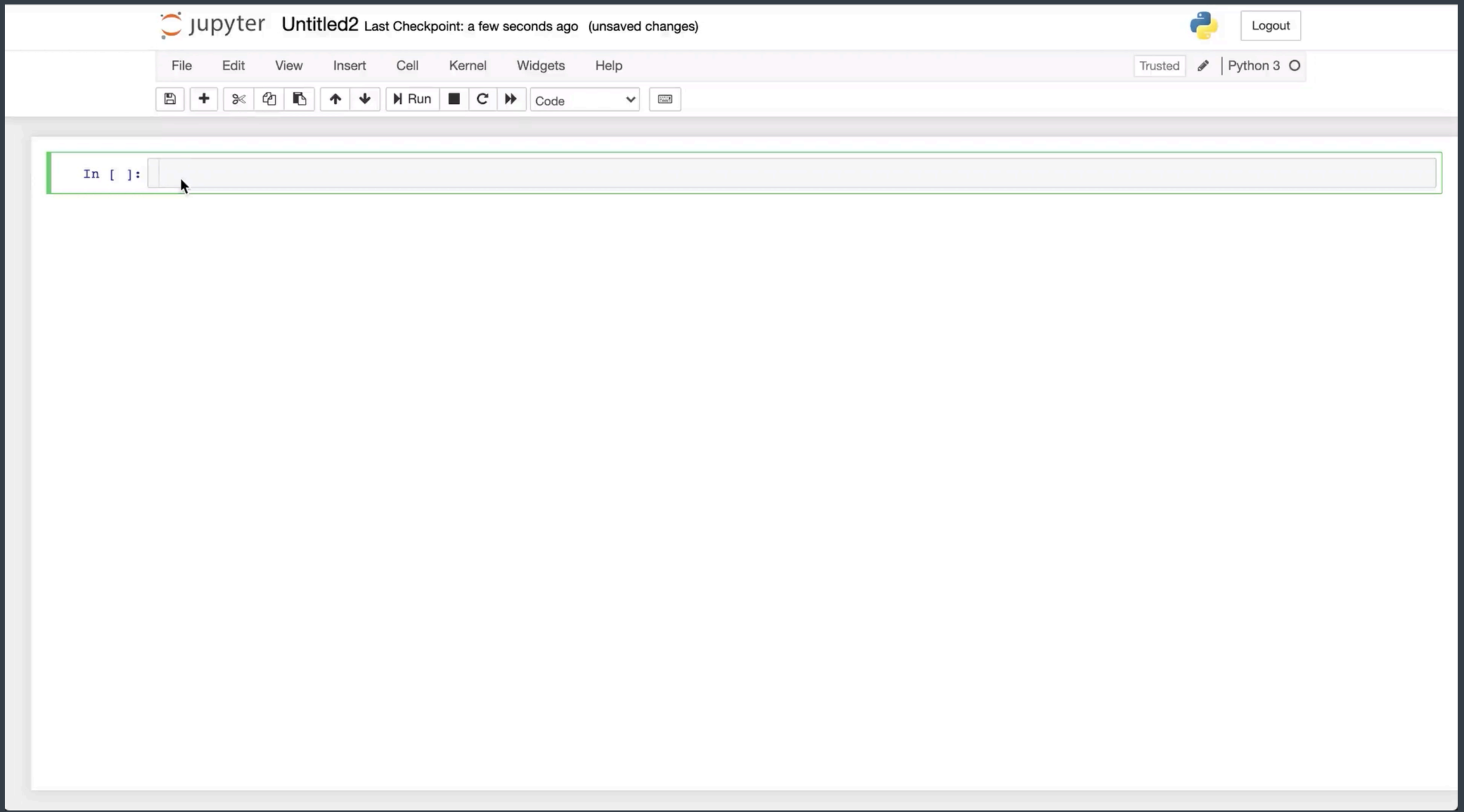
Out[43]: <matplotlib.text.Text at 0x10c9fa490>



[Brian C. Keegan, 2014]

B2: Bridging Code & Interactive Visualization in Notebooks

with



B2: Bridging Code & Interactive Visualization in Notebooks

with

jupyter **Untitled2** Last Checkpoint: a few seconds ago (unsaved changes)

Logout

File Edit View Insert Cell Kernel Widgets Help

Trusted | Python 3

Run Code Toggle Toggle B2 charts Toggle Columns

In [1]:
```
from midas import B2

# init the env
b2 = B2()
```

In [ ]:

Load Data

To load data, use `.from_file("path/to/your_data.csv")`, the columns will show up to the right.

Load Charts

For any dataframes, you can visualize it here with `.vis()`.

Making Interactions

All the loaded charts by default are augmented with interactivity by default. You can **select** a subset of the data by **shift click** with a bar chart, or shift-drag to draw the brush, with a scatter plot or line chart.

Recording/Restoring Interactions

By default, your interactions are executed via a "log" in a code cell to the left. You can look at the code to get a sense of what you have interacted with. You can also execute the code by uncommenting the relevant selections.

Toggle and Resize Panes

To toggle this pane (restoring the traditional notebook view), click on **Toggle midas**, from the menu bar on the top. Similarly, you can click **Toggle Column Shelf** to just hide the pane to the right.
To resize, you can also drag the left edge of the the main (blue)pane---the resizer will be highlighted with a darker shade of blue when you hover over.

Column Pane

Columns of loaded data will show here. You can click on the x symbol to the right to hide the column from view.

See Distributions

You can **click on the column** to get the distribution of the data.

B2: Bridging Code & Interactive Visualization in Notebooks

with


jupyter **Untitled2** Last Checkpoint: a few seconds ago (unsaved changes)

Logout

File Edit View Insert Cell Kernel Widgets Help

Trusted | Python 3

Run | Code | Toggle | Toggle B2 | charts | Toggle Columns

```
# init the env
b2 = B2()
```

In [2]:
```
df = b2.from_file("./data/fire.csv")
df
```

Out[2]:

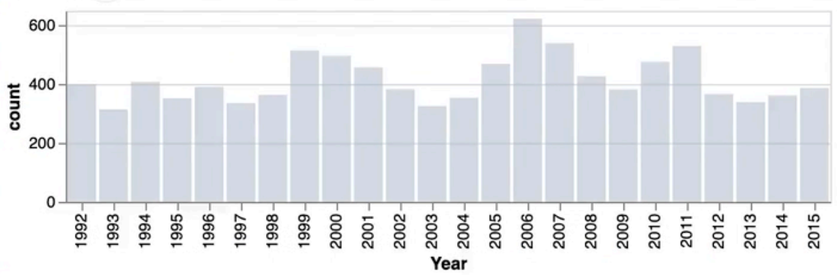
SourceType	Name	Year	DiscoveryDate	DiscoveryTime	Caus
NONFED	nan	1992	2.44865e+06	1540	Arso
NONFED	nan	1998	2.45091e+06	nan	Debris Burnin
NONFED	nan	1995	2.44983e+06	nan	Childre
FED	MINE	2011	2.45575e+06	1800	Lightnin
NONFED	2013039	2013	2.45648e+06	nan	Equipment Us
NONFED	nan	1994	2.44944e+06	1030	Missing/Undefine
NONFED	nan	2015	2.45706e+06	1352	Debris Burnin
NONFED	nan	2000	2.4517e+06	nan	Equipment Us
FED	WAVES AGAIN	2011	2.45575e+06	2144	Campfir
NONFED	MERCED COUNTY	1993	2.44912e+06	nan	Equipment Us

... (9990 rows omitted)

In [4]:
```
Year_dist = df.group('Year')
Year_dist.vis()
```

In [ ]:

Year_dist



df

- SourceType
- Name
- Year
- DiscoveryDate
- DiscoveryTime
- Cause
- ContDate
- ContDoy
- ContTime
- Size
- SizeClass
- Lat
- Lon
- Owner
- State
- County

B2: Bridging Code & Interactive Visualization in Notebooks

with

jupyter **Untitled2** Last Checkpoint: a minute ago (unsaved changes)

Logout

File Edit View Insert Cell Kernel Widgets Help

Trusted | Python 3

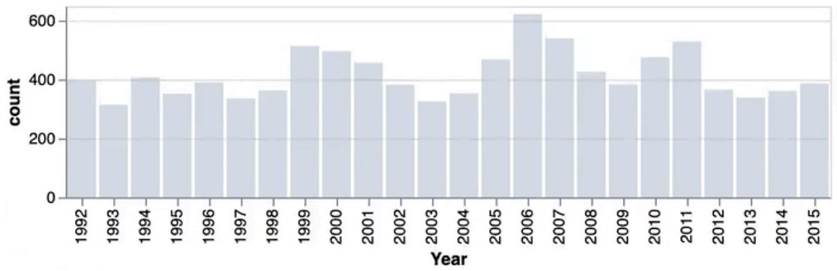
Run | Code | Toggle | Toggle B2 | charts | Toggle Columns

NONFED	nan	1992	2.44865e+06	1540	Arso
NONFED	nan	1998	2.45091e+06	nan	Debris Burnin
NONFED	nan	1995	2.44983e+06	nan	Childre
FED	MINE	2011	2.45575e+06	1800	Lightnin
NONFED	2013039	2013	2.45648e+06	nan	Equipment Us
NONFED	nan	1994	2.44944e+06	1030	Missing/Undefine
NONFED	nan	2015	2.45706e+06	1352	Debris Burnin
NONFED	nan	2000	2.4517e+06	nan	Equipment Us
FED	WAVES AGAIN	2011	2.45575e+06	2144	Campfir
NONFED	MERCED COUNTY	1993	2.44912e+06	nan	Equipment Us

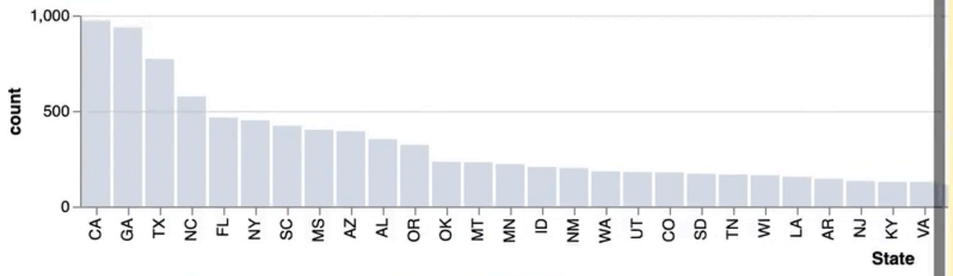
... (9990 rows omitted)

```
In [4]: Year_dist = df.group('Year')
        Year_dist.vis()

In [7]: # 🟡 12:54 PM 🟡
        State_df_dist = df.group('State')
        State_df_dist.vis()

In [8]: # 🟡 12:54 PM 🟡
        Cause_df_dist = df.group('Cause')
        Cause_df_dist.vis()

In [ ]:
```

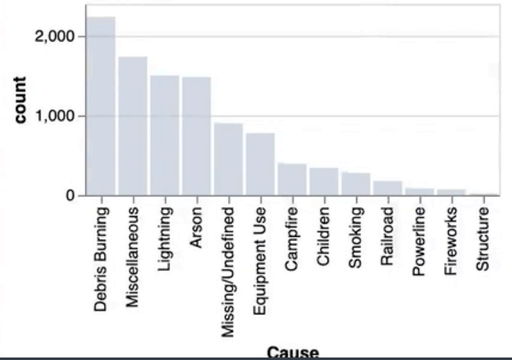
Year_dist



State_df_dist



Cause_df_dist



df

- SourceType
- Name
- Year
- DiscoveryDate
- DiscoveryTime
- Cause
- ContDate
- ContDoy
- ContTime
- Size
- SizeClass
- Lat
- Lon
- Owner
- State
- County

B2: Bridging Code & Interactive Visualization in Notebooks

with

jupyter **Untitled2** Last Checkpoint: a minute ago (unsaved changes)

Logout

File Edit View Insert Cell Kernel Widgets Help

Trusted | Python 3

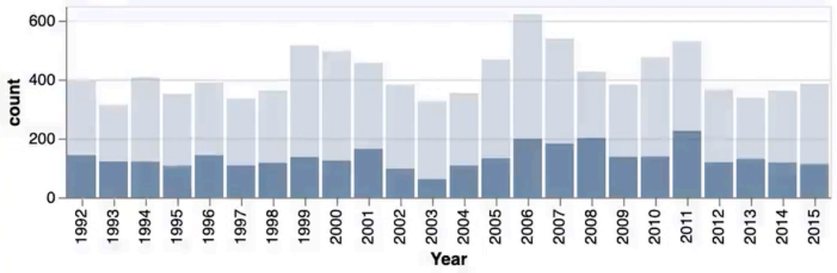
Run | Code | Toggle | Toggle B2 | charts | Toggle Columns

NONFED	nan	1992	2.44865e+06	1540	Arso
NONFED	nan	1998	2.45091e+06	nan	Debris Burnin
NONFED	nan	1995	2.44983e+06	nan	Childre
FED	MINE	2011	2.45575e+06	1800	Lightnin
NONFED	2013039	2013	2.45648e+06	nan	Equipment Us
NONFED	nan	1994	2.44944e+06	1030	Missing/Undefine
NONFED	nan	2015	2.45706e+06	1352	Debris Burnin
NONFED	nan	2000	2.4517e+06	nan	Equipment Us
FED	WAVES AGAIN	2011	2.45575e+06	2144	Campfir
NONFED	MERCED COUNTY	1993	2.44912e+06	nan	Equipment Us

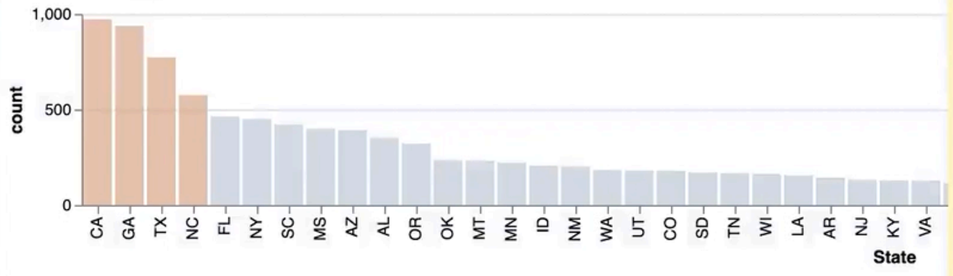
... (9990 rows omitted)

```
In [4]:  Year_dist = df.group('Year')
         Year_dist.vis()

In [7]:  # 🟡 12:54 PM 🟡
         State_df_dist = df.group('State')
         State_df_dist.vis()

In [8]:  # 🟡 12:54 PM 🟡
         Cause_df_dist = df.group('Cause')
         Cause_df_dist.vis()

In [15]: # 🔵 12:55 PM 🔵
         ↔# b2.sel([{"State_df_dist": {"State": ["NC", "TX", "GA
         b2.sel([{"State_df_dist": {"State": ["NC", "TX", "GA", "
```

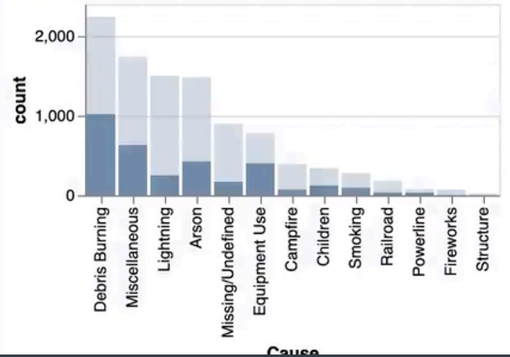
Year_dist



State_df_dist



Cause_df_dist



df

- SourceType
- Name
- Year
- DiscoveryDate
- DiscoveryTime
- Cause
- ContDate
- ContDoy
- ContTime
- Size
- SizeClass
- Lat
- Lon
- Owner
- State
- County

B2: Bridging Code & Interactive Visualization in Notebooks

with


jupyter **Untitled2** Last Checkpoint: a minute ago (unsaved changes)

Logout

File Edit View Insert Cell Kernel Widgets Help

Trusted | Python 3

Run | Code | Toggle | Toggle B2 | charts | Toggle Columns

NONFED	2013039	2013	2.45648e+06	nan	Equipment Us
NONFED	nan	1994	2.44944e+06	1030	Missing/Undefine
NONFED	nan	2015	2.45706e+06	1352	Debris Burnin
NONFED	nan	2000	2.4517e+06	nan	Equipment Us
FED	WAVES AGAIN	2011	2.45575e+06	2144	Campfir
NONFED	MERCED COUNTY	1993	2.44912e+06	nan	Equipment Us

... (9990 rows omitted)

In [4]:
```
Year_dist = df.group('Year')
Year_dist.vis()
```

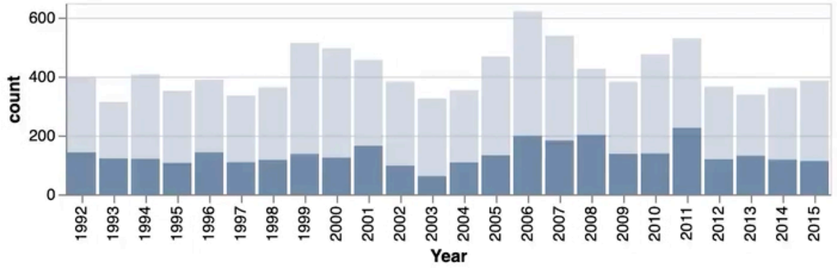
In [7]:
```
# 🟡 12:54 PM 🟡
State_df_dist = df.group('State')
State_df_dist.vis()
```

In [8]:
```
# 🟡 12:54 PM 🟡
Cause_df_dist = df.group('Cause')
Cause_df_dist.vis()
```

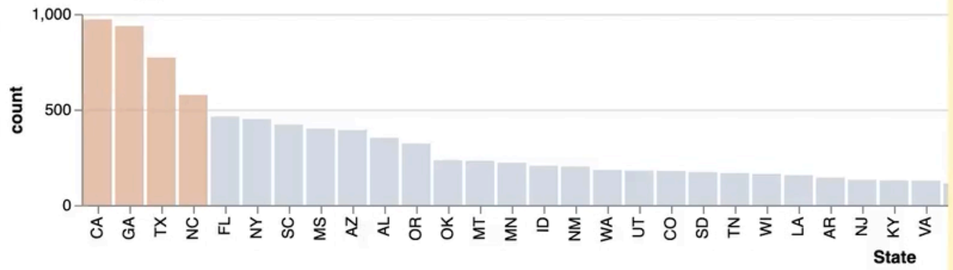
In [18]:
```
# 🔵 12:55 PM 🔵
# b2.sel([{"State_df_dist": {"State": ["CA"]}}])
# b2.sel([{"State_df_dist": {"State": ["GA"]}}])
# b2.sel([{"State_df_dist": {"State": ["TX"]}}])
# b2.sel([{"State_df_dist": {"State": ["NC"]}}])
# b2.sel([{"State_df_dist": {"State": ["NC", "TX"]}}])
# b2.sel([{"State_df_dist": {"State": ["NC", "TX", "GA"]
b2.sel([{"State_df_dist": {"State": ["NC", "TX", "GA", "
```

In [ ]:

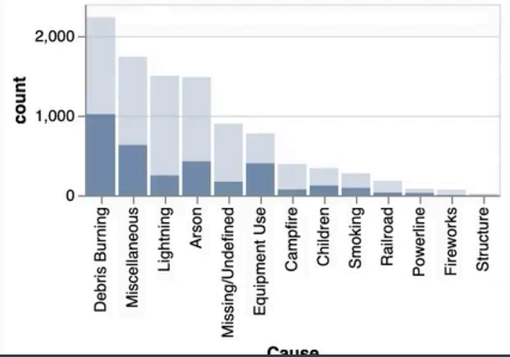
Year_dist



State_df_dist



Cause_df_dist



df

- SourceType
- Name
- Year
- DiscoveryDate
- DiscoveryTime
- Cause
- ContDate
- ContDoy
- ContTime
- Size
- SizeClass
- Lat
- Lon
- Owner
- State
- County

B2: Bridging Code & Interactive Visualization in Notebooks

with


jupyter **Untitled2** Last Checkpoint: 17 minutes ago (autosaved)

Logout

File Edit View Insert Cell Kernel Widgets Help

Trusted | Python 3

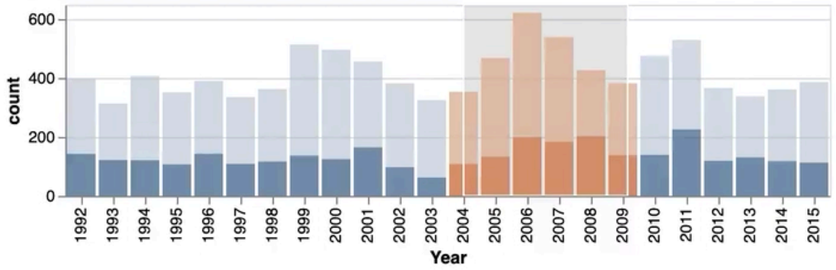
Run | Code | Toggle | Toggle B2 | charts | Toggle Columns

NONFED	nan	2000	2.4517e+06	nan	Equipment Us
FED	WAVES AGAIN	2011	2.45575e+06	2144	Campfir
NONFED	MERCED COUNTY	1993	2.44912e+06	nan	Equipment Us

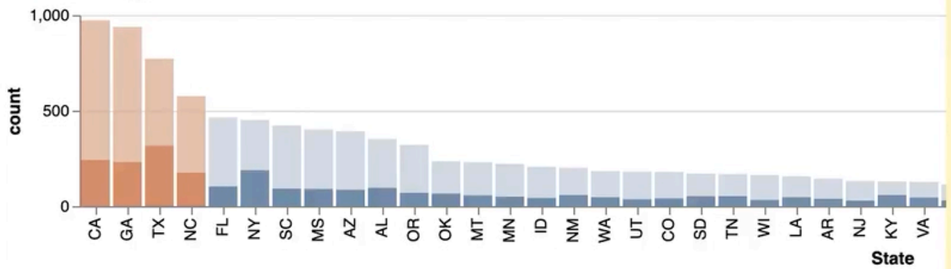
... (9990 rows omitted)

```
In [4]:  Year_dist = df.group('Year')
         Year_dist.vis()

In [7]:  # 🟡 12:54 PM 🟡
         State_df_dist = df.group('State')
         State_df_dist.vis()

In [8]:  # 🟡 12:54 PM 🟡
         Cause_df_dist = df.group('Cause')
         Cause_df_dist.vis()

In [35]: # 🔵 12:55 PM 🔵
         # b2.sel([{"State_df_dist": {"State": ["CA"]}}])
         # b2.sel([{"State_df_dist": {"State": ["GA"]}}])
         # b2.sel([{"State_df_dist": {"State": ["TX"]}}])
         # b2.sel([{"State_df_dist": {"State": ["NC"]}}])
         # b2.sel([{"State_df_dist": {"State": ["NC", "TX"]}}])
         # b2.sel([{"State_df_dist": {"State": ["NC", "TX", "GA"]
         b2.sel([{"State_df_dist": {"State": ["NC", "TX", "GA", "

In [37]:
         dist": {"Year": [1998, 1999, 2000, 2001, 2002, 2003]}}])
         st": {"Year": [2004, 2005, 2006, 2007, 2008, 2009]}}])

In [ ]:
```

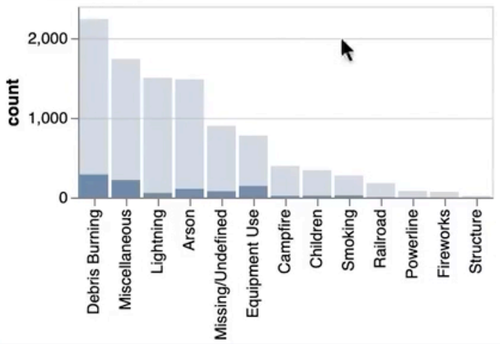
Year_dist



State_df_dist



Cause_df_dist



df

- SourceType
- Name
- Year
- DiscoveryDate
- DiscoveryTime
- Cause
- ContDate
- ContDoy
- ContTime
- Size
- SizeClass
- Lat
- Lon
- Owner
- State
- County

B2: Bridging Code & Interactive Visualization in Notebooks

with


jupyter **Untitled2** Last Checkpoint: 22 minutes ago (unsaved changes)

Logout

File Edit View Insert Cell Kernel Widgets Help

Trusted | Python 3

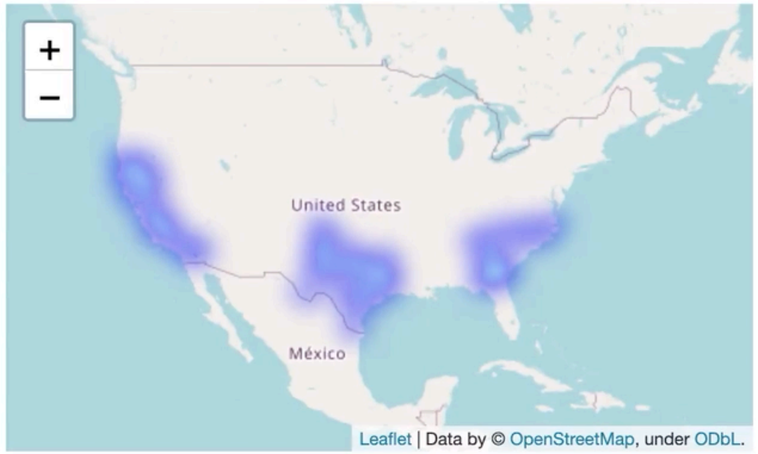
Run Code Toggle Toggle B2 charts Toggle Columns

Children	22
Debris Burning	286
Equipment Use	143
Fireworks	2
Lightning	56
Miscellaneous	219
Missing/Undefined	77
Powerline	2

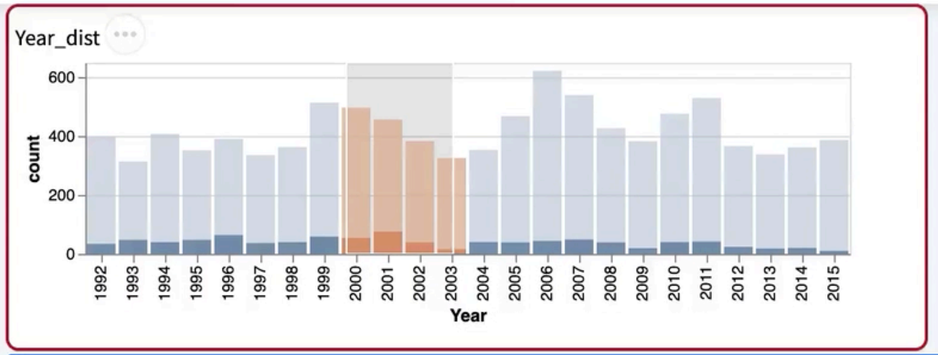
... (2 rows omitted)

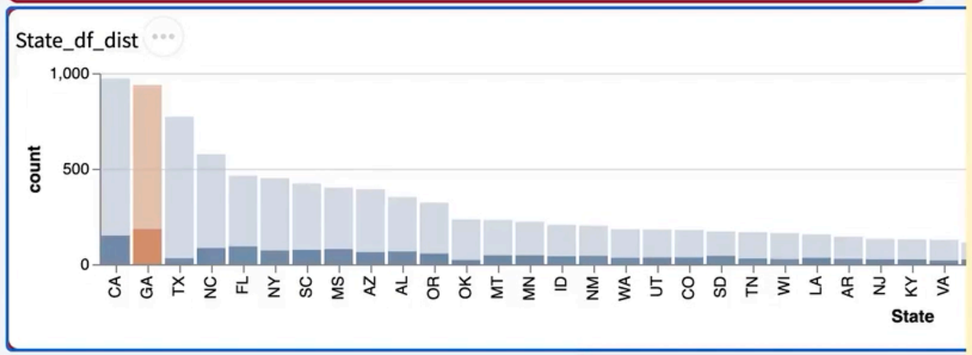
In [40]:
```
locs = df.where('State', b2.are.contained_in(['NC', 'TX'
locs.plot_heatmap(zoom_start=3, radius=6)
```

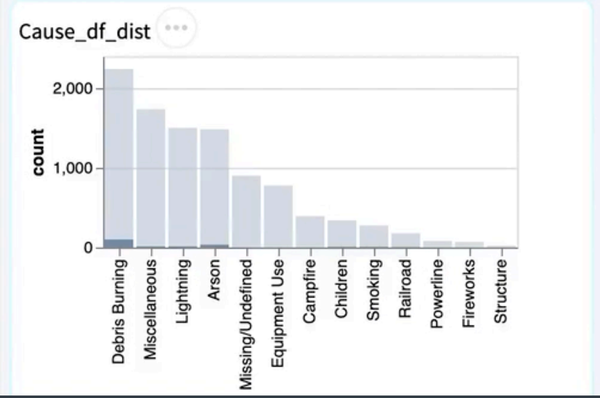
Out[40]:



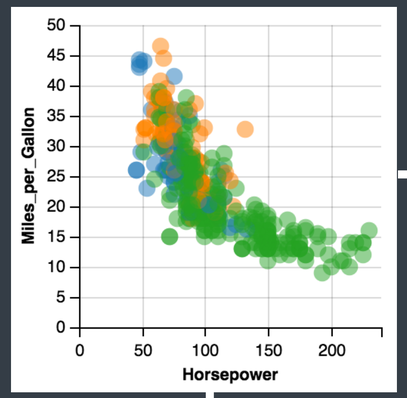
In [42]:
```
# 🔵 01:16 PM 🔵
# b2.sel([{"State_df_dist": {"State": ["NC", "TX", "GA",
b2.sel([{"State_df_dist": {"State": ["NC", "TX", "GA", "
```



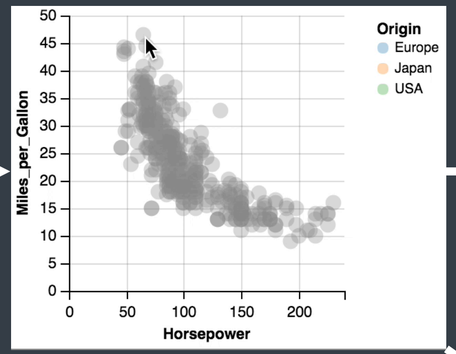




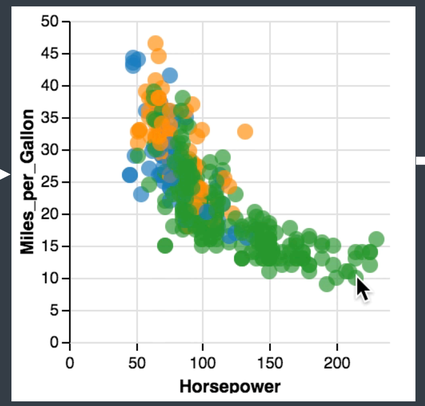
df
- SourceType
- Name
- Year
- DiscoveryDate
- DiscoveryTime
- Cause
- ContDate
- ContDoy
- ContTime
- Size
- SizeClass
- Lat
- Lon
- Owner
- State
- County



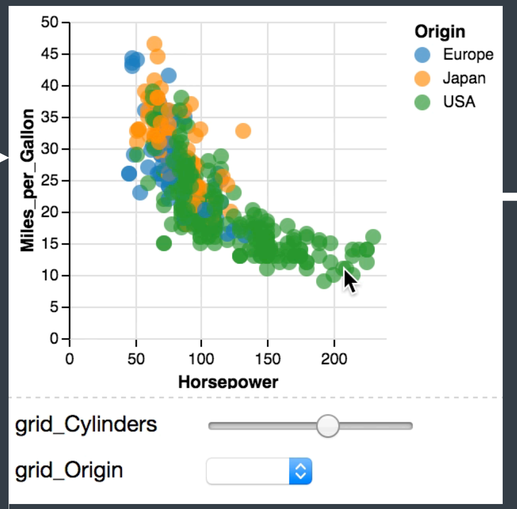
Single



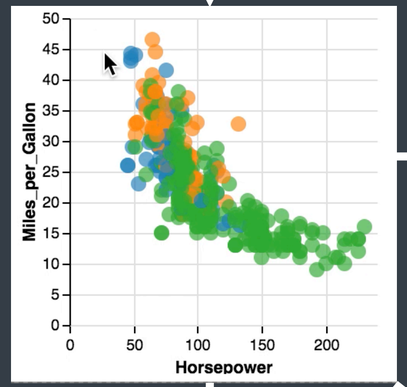
Project



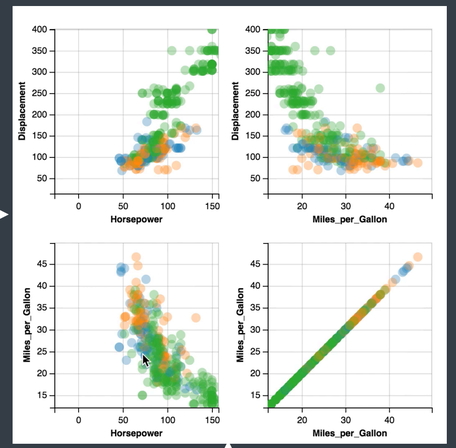
Bind



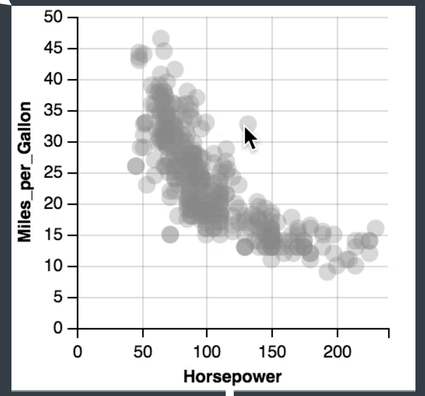
Interval



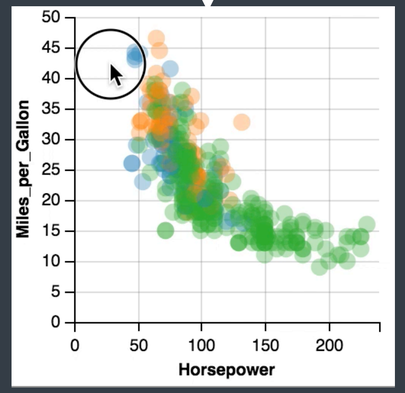
Repeat



Multi

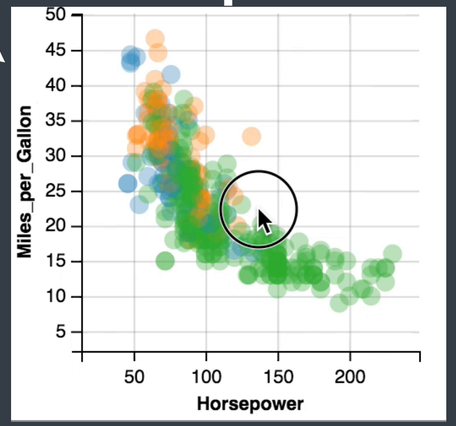


Project

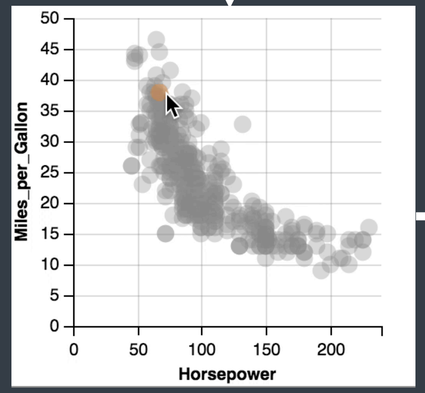


Bind

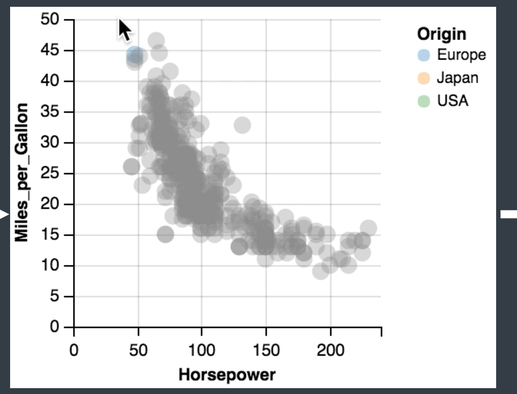
Repeat



Hover



Nearest



Systematically Enumerate the Design Space

Lyra: Interactive Visualization by Demonstration

with



GROUPS +

Edit Scene

Group 1

GUIDES

MARKS

INTERACTIONS

WIDGETS

SCALES

GROUP 1

PIPELINE None

X POSITION

Left 0

Set to group width

Width 640

Y POSITION

Top 0

Set to group height

Height 360

FILL

Color transparent

Opacity 1

STROKE

Color

Width 0

DATA PIPELINES World + New

Birth_Rate	0.011	0.019	0.019
Life_Exp...	77	78	77
Life_Exp...	70	68	73
Country/...	Mauritius	Seychelles	Tunisia
GDP_per...	8862	11689	4197

Region	Africa	Africa	Africa
Infant_M...	0.013	0.012	0.014
Year	135433...	135433...	135433...
Life_Exp...	74	73	75

1–20 of 207

RECT

SYMBOL

TEXT

LINE

AREA

EXPORT

Lyra: Interactive Visualization by Demonstration

with



GROUPS +

Edit Scene

Group 1

GUIDES

MARKS

Symbol 1

INTERACTIONS

WIDGETS

SCALES

SYMBOL 1

PIPELINE None

POSITION

X 100

Y 100

GEOMETRY

Shape circle

Size 200

FILL

Color #4682b4

Opacity 1

STROKE

Color #000000

Width 0

DATA PIPELINES World + New

Birth_Rate	0.011	0.019	0.019
# ··· Life_Expectancy_Female			
Life_Exp...	70	68	73
Country/...	Mauritius	Seychelles	Tunisia
GDP_per...	8862	11689	4197

Region	Africa	Africa	Africa
Infant_M...	0.013	0.012	0.014
Year	135433...	135433...	135433...
Life_Exp...	74	73	75

··· Life_Expectancy_Female

1–20 of 20

RECT

SYMBOL

TEXT

LINE

AREA

EXPORT

Lyra: Interactive Visualization by Demonstration

with



GROUPS +

Edit Scene

Group 1

GUIDES

MARKS

Symbol 1

INTERACTIONS

WIDGETS

SCALES

SYMBOL 1

PIPELINE None

POSITION

X 100

Y 100

GEOMETRY

Shape circle

Size 200

FILL

Color #4682b4

Opacity 1

STROKE

Color #000000

Width 0

DATA PIPELINES World + New

Birth_Rate	0.011	0.019	0.019
# ··· Life_Expectancy_Female			
Life_Exp...	70	68	73
Country/...	Mauritius	Seychelles	Tunisia
GDP_per...	8862	11689	4197

Region	Africa	Africa	Africa
Infant_M...	0.013	0.012	0.014
Year	135433...	135433...	135433...
Life_Exp...	74	73	75

1–20 of 207

RECT

SYMBOL

TEXT

LINE

AREA

EXPORT

Lyra: Interactive Visualization by Demonstration

with



GROUPS +

Edit Scene

▼ Group 1

GUIDES

X Axis

Y Axis

Color Legend

MARKS

Symbol 1

INTERACTIONS +

WIDGETS

SCALES

SYMBOL 1

PIPELINE World

POSITION

X x Life_Expec...

Y y Birth_Rate

GEOMETRY

Shape circle

Size 200

FILL

Color color Region

Opacity 1

STROKE

Color #000000

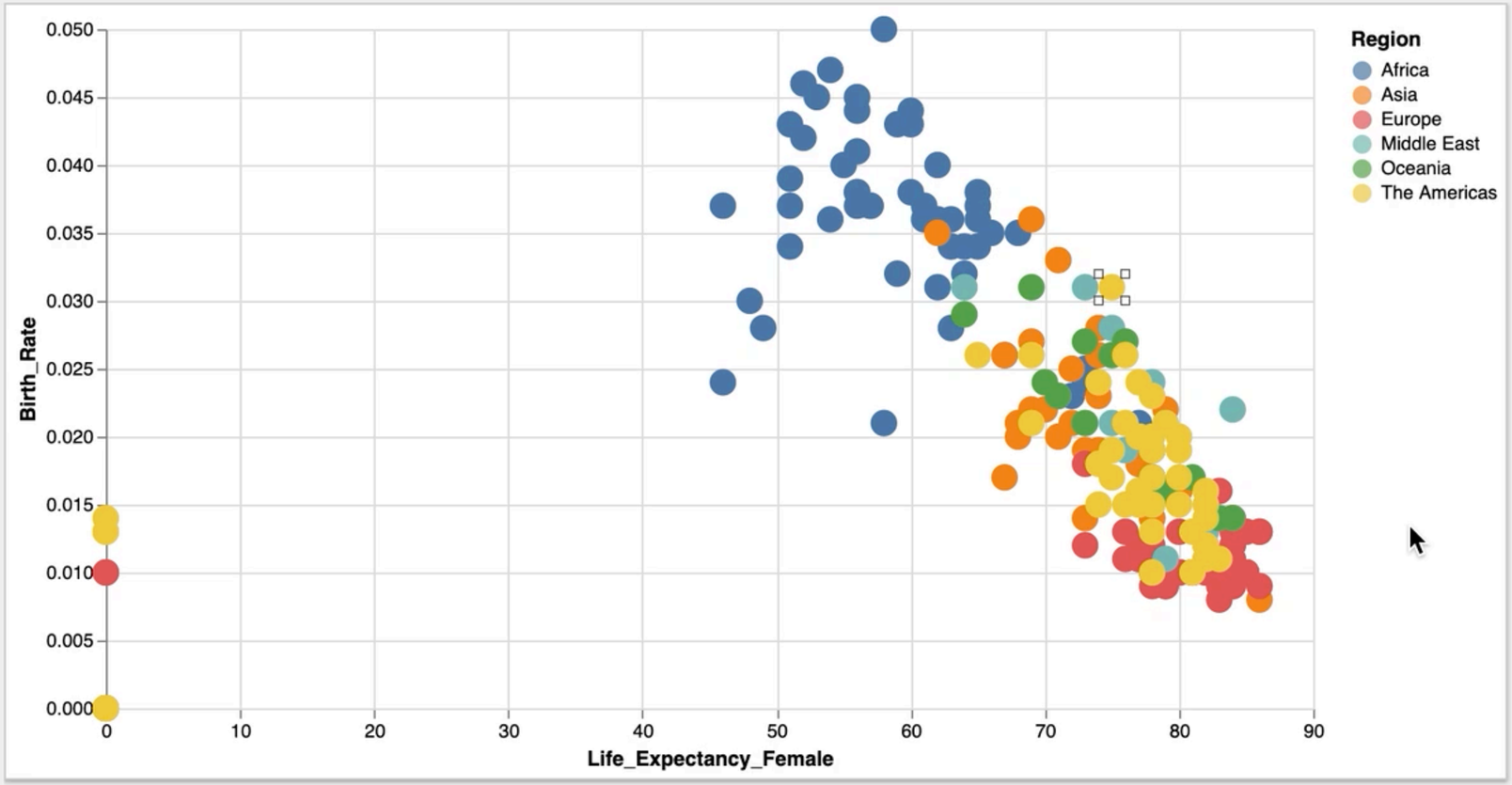
Width 0

DATA PIPELINES World + New

Birth_Rate	0.011	0.019	0.019
Life_Exp...	77	78	77
Life_Exp...	70	68	73
Country/...	Mauritius	Seychelles	Tunisia
GDP_per...	8862	11689	4197

Region	Africa	Africa	Africa
Infant_M...	0.013	0.012	0.014
Year	135433...	135433...	135433...
Life_Exp...	74	73	75

1–20 of 207



RECT

SYMBOL

TEXT

LINE

AREA

INTERACTION

EXPORT

Lyra: Interactive Visualization by Demonstration

with



GROUPS +

Edit Scene

▼ Group 1

GUIDES

X Axis

Y Axis

Color Legend

MARKS

Symbol 1

INTERACTIONS +

Interaction 1

WIDGETS

SCALES

INTERACTION 1

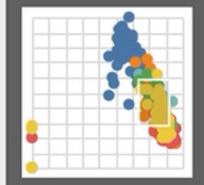
INPUT EVENT

Mouse Drag

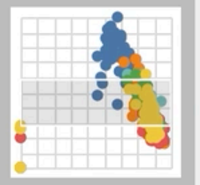
Keyboard None

SELECTIONS

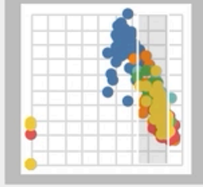
Brush



Brush (y-axis)

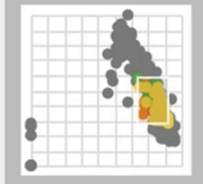


Brush (x-axis)

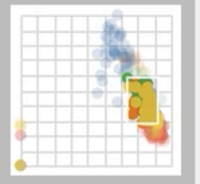


APPLICATIONS

Color



Opacity



Size

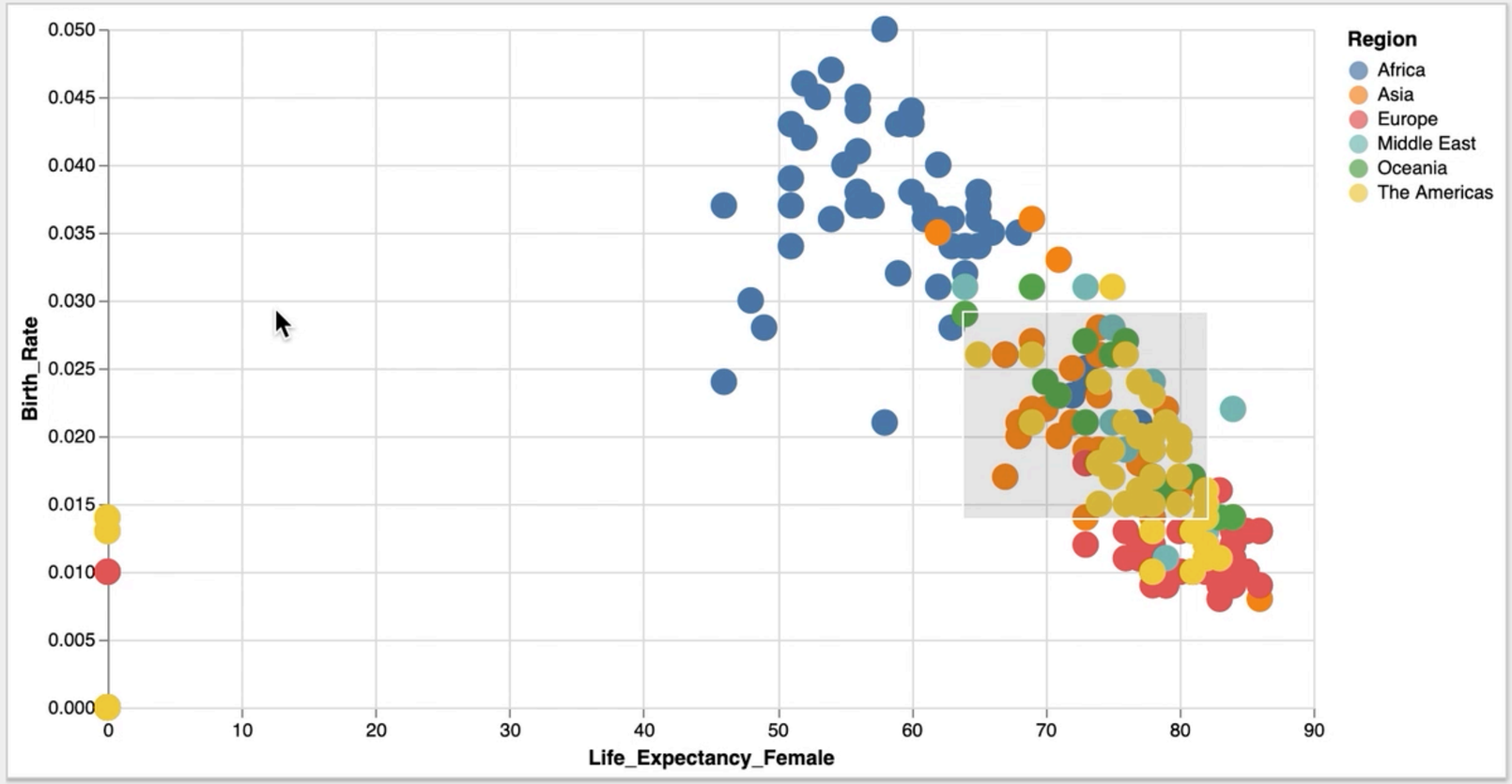
Pan and zoom

DATA PIPELINES World + New

Birth_Rate	0.011	0.019	0.019
Life_Exp...	77	78	77
Life_Exp...	70	68	73
Country/...	Mauritius	Seychelles	Tunisia
GDP_per...	8862	11689	4197

Region	Africa	Africa	Africa
Infant_M...	0.013	0.012	0.014
Year	135433...	135433...	135433...
Life_Exp...	74	73	75

1–20 of 207



RECT

SYMBOL

TEXT

LINE

AREA

INTERACTION

EXPORT

Towards *Effective* Interaction With Data Visualization

Empirically Derive Effectiveness "Rankings"

Use an interaction grammar to enumerate alternate techniques for a given analytic task. Test with human subjects.

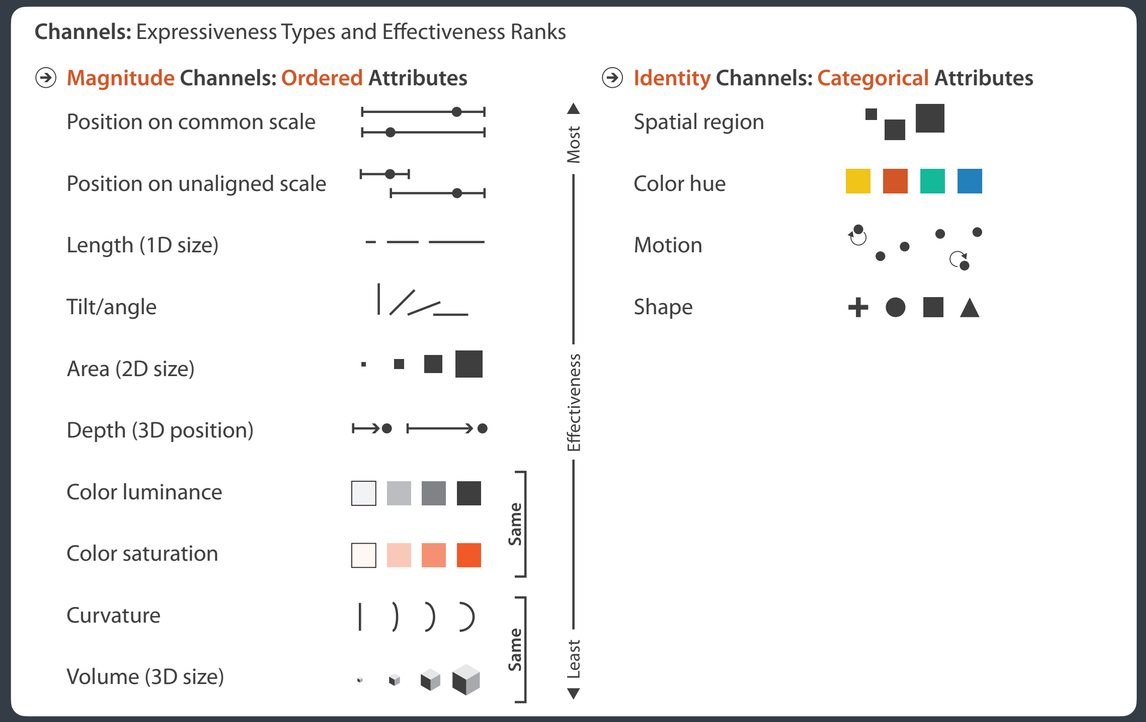
A key challenge: measuring effectiveness beyond time/error?

Interaction Recommender Systems

Synthesize appropriate interactive visualizations using effectiveness rankings.

Mine interaction histories recorded as semantically-meaningful *selections*. Accelerate interactive analysis, suggest unexplored paths.

A key challenge: how to infer user intent from context?



Towards *Effective* Interaction With Data Visualization

Empirically Derive Effectiveness "Rankings"

Use an interaction grammar to enumerate alternate techniques for a given analytic task. Test with human subjects.

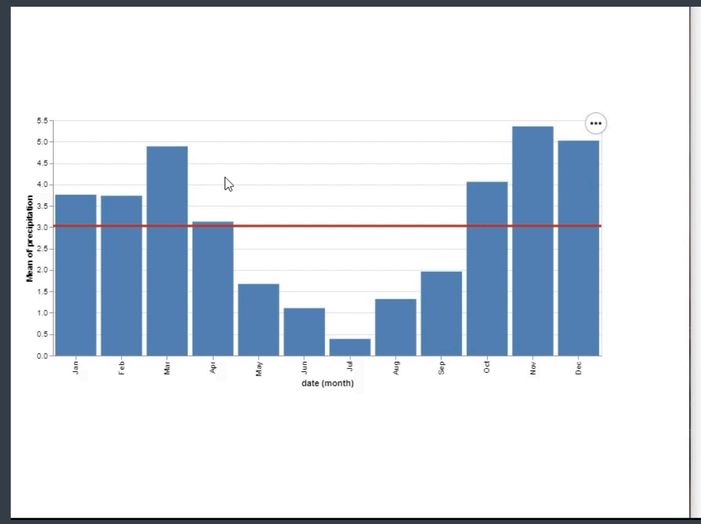
A key challenge: measuring effectiveness beyond time/error?

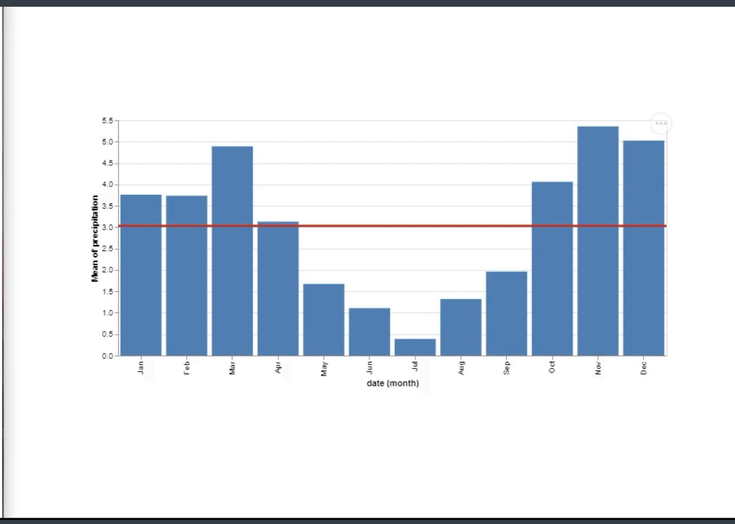
Interaction Recommender Systems

Synthesize appropriate interactive visualizations using effectiveness rankings.

Mine interaction histories recorded as semantically-meaningful *selections*. Accelerate interactive analysis, suggest unexplored paths.

A key challenge: how to infer user intent from context?





Supporting (A)Synchronous Collaboration

Establishing common ground and shared awareness.

A key challenge: interactive visualization beyond the desktop?

Towards *Effective* Interaction

With Data Visualization

Arvind Satyanarayan
@arvindsatya1

MIT Visualization Group
@mitvis • vis.csail.mit.edu

+ friends