

## Lesson 4

# Visualising and Analysing High-dimensional Data

Instructor: Dr. Kam Tin Seong  
Associate Professor of Information System (Practice)  
School of Information Systems  
Singapore Management University

### What will you learn from this lesson?

---

- Sensing and analyzing high-dimensional continuous data
- Sensing and analyzing high-dimensional categorical data
- Sensing both categorical and continuous high-dimensional data
- High-dimensional data analysis best practices

## Beyond bivariate data

- Three continuous variables

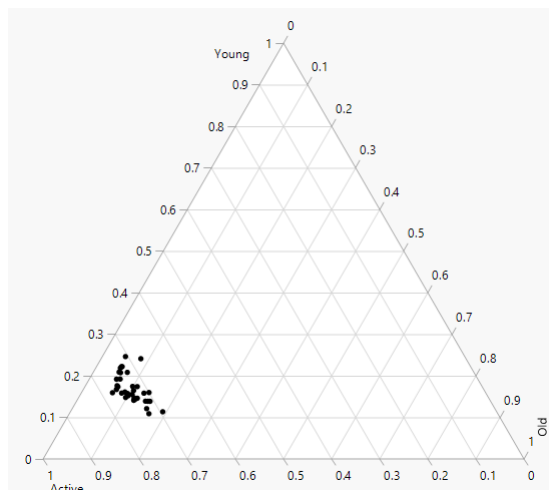
CENSUS OF POPULATION 2010  
 Table A3 Resident Population by Planning Area, Age Group and Sex

Planning Area	Total	Number													
		0-4	5-9	10-14	15-19	20-24	25-29	30-34	35-39	40-44	45-49	50-54	55-59	60-64	65 & Over
Total	3,771,721	194,432	215,675	244,302	263,750	247,190	272,639	298,687	320,024	309,441	323,459	303,044	248,696	191,995	338,387
Ang Mo Kio	179,297	7,967	8,424	9,335	10,457	10,656	13,400	14,502	14,510	13,525	14,862	14,605	13,785	11,868	21,401
Bedok	294,519	13,230	15,018	17,489	20,083	20,156	21,265	21,707	22,751	22,018	24,615	24,632	21,891	18,018	31,646
Bishan	91,298	3,941	4,759	5,691	7,188	6,338	5,930	6,010	7,354	7,102	8,041	8,245	6,725	5,141	8,833
Bukit Batok	144,198	7,187	8,516	9,738	10,427	10,625	11,332	10,941	11,829	12,259	13,049	12,433	9,911	6,487	9,464
Bukit Merah	157,122	8,049	6,892	6,894	7,479	7,940	10,865	13,871	13,495	11,796	11,650	12,057	11,352	10,782	24,000
Bukit Panjang	128,734	7,166	8,334	9,324	10,224	9,734	9,102	9,488	11,171	10,469	11,528	10,784	7,907	5,316	8,247
Bukit Timah	70,314	3,115	4,416	4,681	3,238	4,613	3,816	4,096	5,889	6,071	6,156	5,726	4,841	4,378	7,638
Changi	2,155	169	192	159	139	94	111	187	252	223	175	122	81	73	178
Choa Chu Kang	173,291	9,232	11,823	15,195	15,362	11,637	11,893	12,541	14,483	16,675	16,919	13,299	9,112	5,855	9,265
Clementi	91,874	4,235	4,567	4,713	4,732	5,133	7,093	7,818	7,831	7,201	7,331	6,882	7,064	6,648	10,626
Downtown Core	3,722	144	122	140	137	161	287	315	326	357	295	295	265	262	616



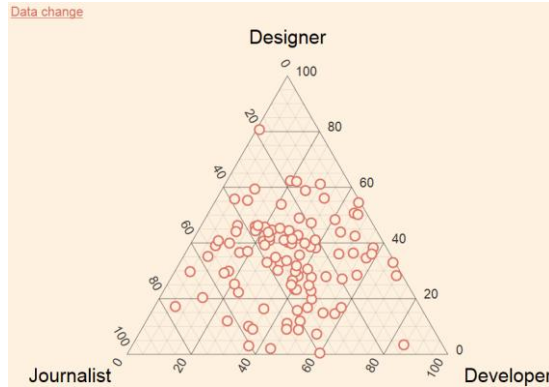
Planning Zone	Young	Active	Old
Outram	2260	13768	3811
Downtown Core	406	2700	616
Rochor	1899	11341	2424
Queenstown	13777	69670	15055
Bukit Merah	21835	111287	24000
Toa Payoh	17327	80780	18538
Kallang	13861	71438	14260
Marine Parade	7573	33116	8629
Novena	7389	33164	6067
Geylang	17657	88237	14796
Ang Mo Kio	25726	132170	21401
Others	634	3322	528
Clementi	13515	67733	10626
Bukit Timah	12252	50424	7638
Bedok	45737	217136	31646

## Ternary chart



### D3.js based Ternary Chart

- Ternary Plot in D3.js (<http://bl.ocks.org/tomgp/7674234>), (<https://gist.github.com/widged/5780720>), (<http://bl.ocks.org/tomgp/7766353>)

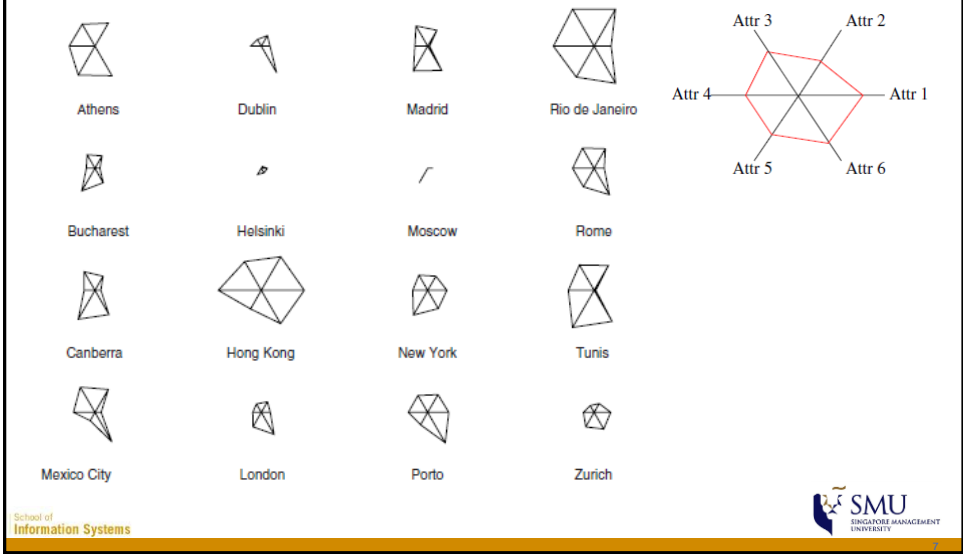


### Multivariate data: climatic variables

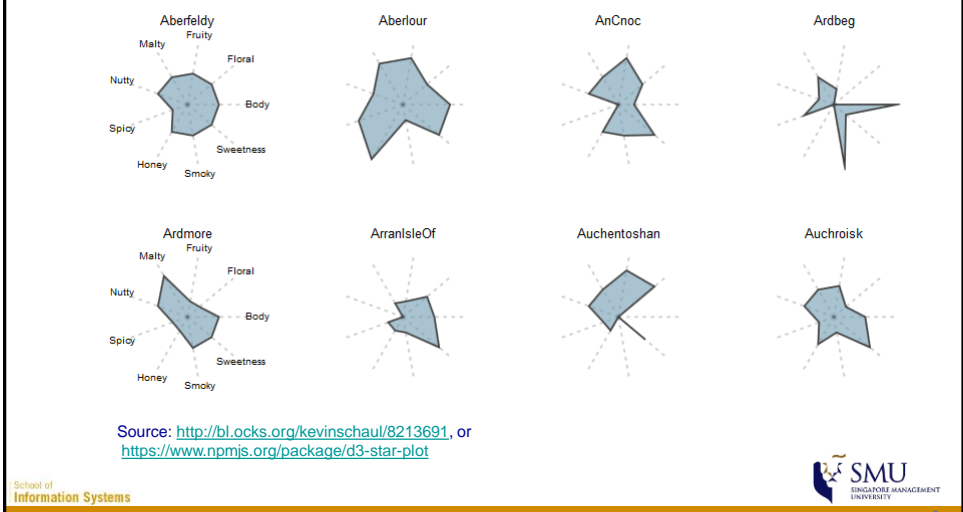
City	Precip. average	Temp. average	Temp. max average	Temp. min average	Record max	Record min
Athens	37	17	21	13	42	-3
Bucharest	58	11	16	5	49	-23
Canberra	62	12	19	6	42	-10
Dublin	74	10	12	6	28	-7
Helsinki	63	5	8	1	31	-36
Hong Kong	218	23	25	21	37	2
London	75	10	13	5	35	-13
Madrid	45	13	20	7	40	-10
Mexico City	63	17	23	11	32	-3
Moscow	59	4	8	1	35	-42
New York	118	12	17	8	40	-18
Porto	126	14	18	10	34	-2
Rio de Janeiro	109	25	30	20	43	7
Rome	80	15	20	11	37	-7
Tunis	44	18	23	13	46	-1
Zurich	107	9	12	6	35	-20

Table 4.1 Annual climatic values in Celsius of some world cities. Values from <http://www.weatherbase.com>.

## Star Plot



## Static Star Plot in D3.js



### Interactive Radar Chart

Source: <http://graves.cl/radar-chart-d3/> and <http://bl.ocks.org/nbremer/6506614>

School of Information Systems

### Radar Chart in D3.js Reusable Library

Source: <https://github.com/alangrafu/radar-chart-d3>

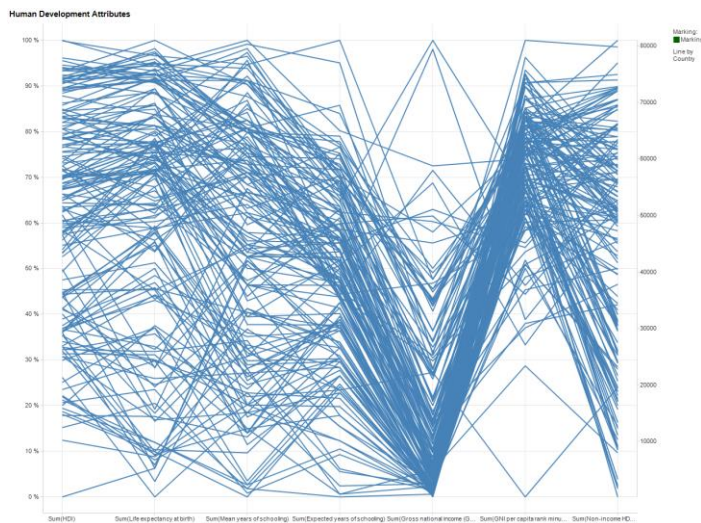
School of Information Systems

## High-dimensional continuous data

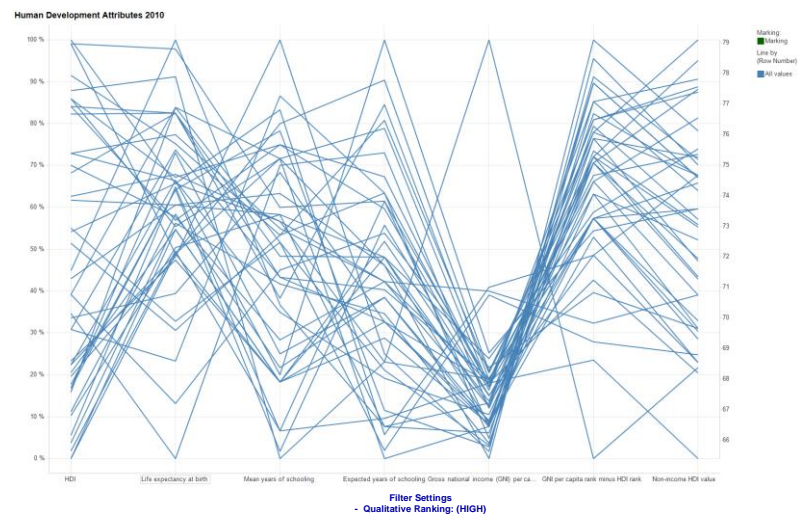
- Multivariate data
- More than three variables

ADM2NAME	ADM2 CODE	Entry cost	Land access	Transparency	Time Costs	Informal charges	Proactivity	Business support...	Labor training	Legal institutions	unweighted PCI	PCI 2010
Ha Noi	101	6.079872	3.036629	5.622983	5.462117	5.695304	3.200226	7.600075	6.724179	5.097846	48.519231	55.7301765
Hai Phong	103	6.378953	3.831609	6.232802	5.61721	6.078024	2.655746	6.959484	5.467642	4.603246	47.824716	54.6365965
Vinh Phuc	104	6.601418	6.019302	5.606139	6.906237	5.843651	8.082954	5.166162	5.695219	5.289262	55.210344	61.7264575
Bac Ninh	106	7.291559	5.42377	6.365728	7.693029	5.884039	7.086143	5.810097	5.761192	5.638832	56.946389	64.478474
Hai Duong	107	6.508549	5.94212	5.374836	6.684187	6.244188	5.055763	6.216137	5.269055	4.624822	51.919657	57.514102
Hung Yen	109	5.806461	6.136844	5.420558	7.467897	5.295288	3.267085	4.660326	2.955754	4.089105	45.101318	49.768441
Ha Nam	111	5.982533	5.374191	5.562412	4.806478	5.549887	4.082179	5.255798	5.350751	4.422565	46.388794	52.176919
Nam Dinh	113	5.659831	6.306973	5.282395	6.883652	6.745993	4.137132	5.989564	5.296639	2.917283	49.219462	55.633412
Thai Binh	115	6.538374	6.07514	5.877275	7.425593	6.091346	5.013376	6.919116	5.350055	4.611089	53.901364	60.039818
Ninh Binh	117	6.768809	5.57222	5.609146	8.311043	6.203602	5.493979	6.979725	5.79477	5.65709	56.390384	62.845304
Ha Giang	201	7.03553	5.501739	5.061701	4.54352	5.189184	6.787162	6.116573	5.143668	3.787485	49.166562	53.9407925
Cao Bang	203	7.468838	4.510241	5.173515	5.534895	5.829948	4.659263	4.618188	5.513456	2.70909	46.017434	53.553093
Lao Cai	205	7.707034	7.463224	7.391482	7.269845	7.157474	6.941723	6.318152	5.708106	4.294513	60.251553	67.948119
Bac Can	207	6.283001	6.33004	5.27312	5.605261	4.855072	4.551651	5.226332	4.653521	3.514567	46.292565	51.486367
Lang Son	209	6.21592	4.204638	5.410324	4.693871	5.98648	3.619392	6.842519	4.861953	2.544278	44.379375	50.20287
Tuyen Quang	211	5.218991	5.191109	6.857085	5.819696	6.041488	5.97811	5.356539	5.479868	3.964347	49.907131	57.8979325
Yen Bai	213	6.4728	6.612154	6.012499	7.057146	6.517362	5.501385	6.540938	4.974207	5.065648	54.754137	60.160047
Thai Nguyen	215	5.979152	6.246058	5.425032	6.747609	6.651042	4.781002	5.186144	5.126515	4.380002	50.522156	56.5419055
Phu Tho	217	6.541923	4.335423	5.491023	5.735717	6.842435	3.961593	5.211816	4.319536	4.681257	48.900523	52.4747925
Bac Giang	221	6.443639	4.798666	6.108578	5.831342	6.429758	5.498727	6.296986	5.357154	4.846302	51.611152	58.021578

## Sensing High-dimensional Data with Parallel Coordinates



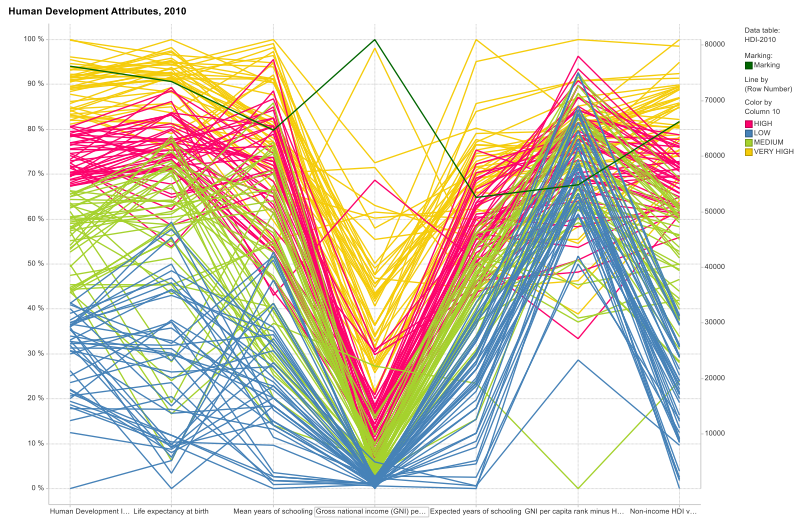
### Explore parallel coordinates with interactive filtering technique



School of Information Systems



### Explore parallel coordinates with pre-attentive property

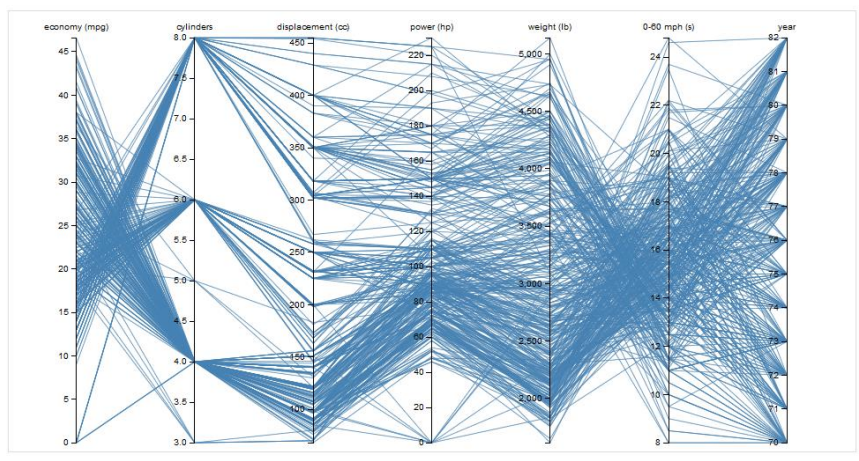


School of Information Systems





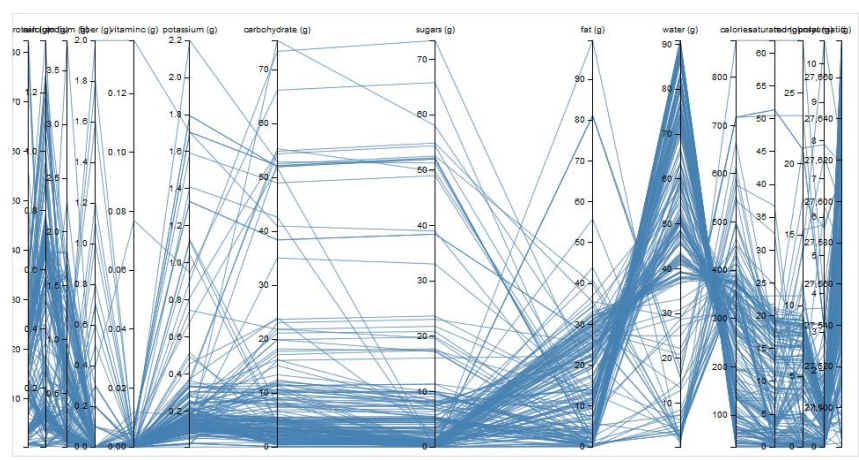
### Parallel Coordinates in d3.js



School of Information Systems



### Parallel Coordinates with Fisheye

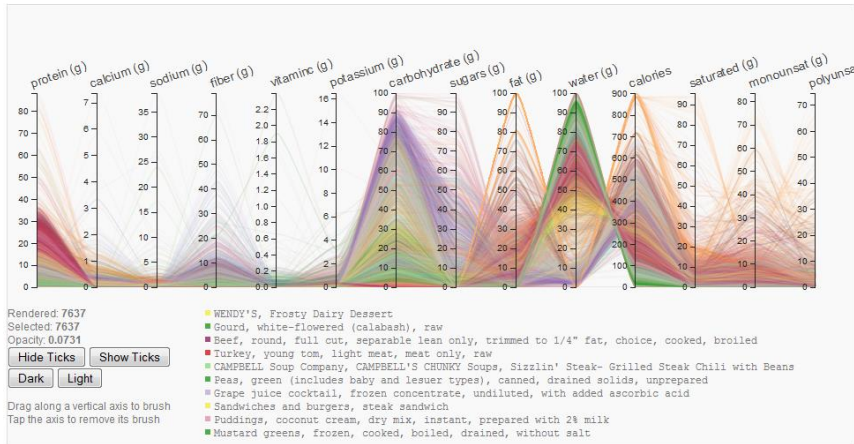


School of Information Systems



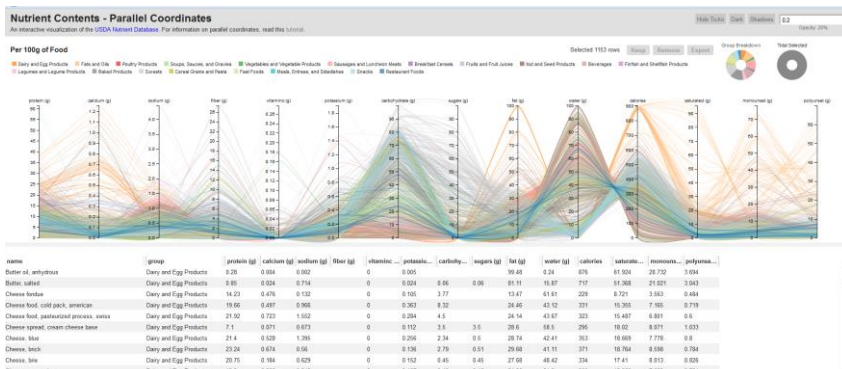


## Parallel Coordinates - Shuffled Rendering

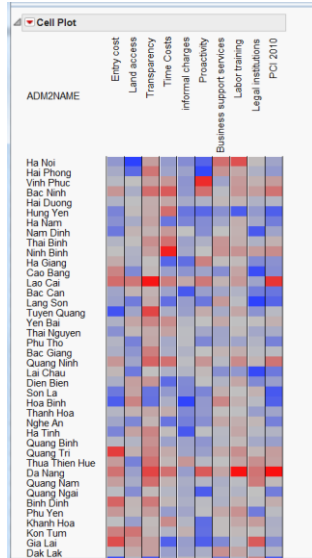


## Data Visualisation Application

- Parallel Coordinates



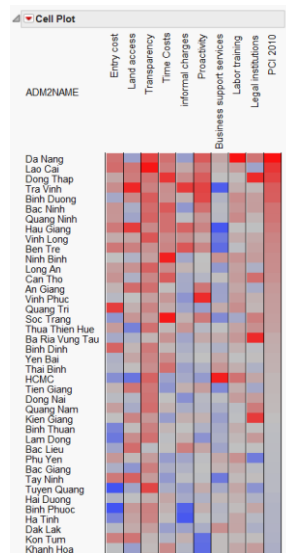
## Sensing High-dimensional Data with Heatmap



School of Information Systems



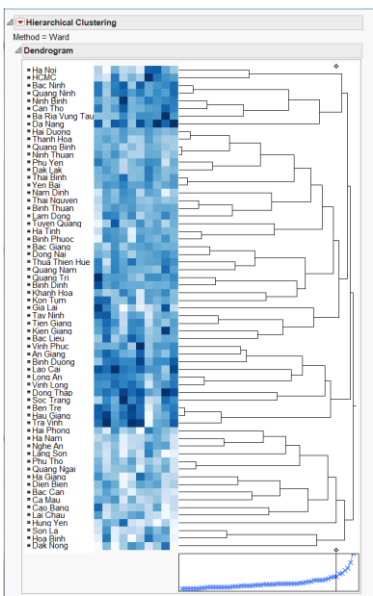
## Heatmap: Sorting



School of Information Systems



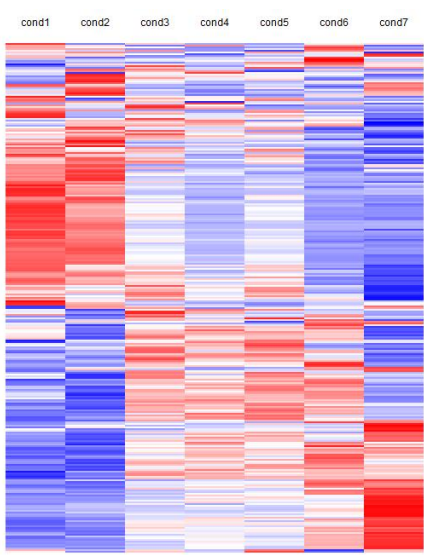
### Heatmap for Exploratory Modelling Analysis



School of Information Systems



### Heatmap library in D3.js



Source: <http://blog.nextgenetics.net/?e=44>

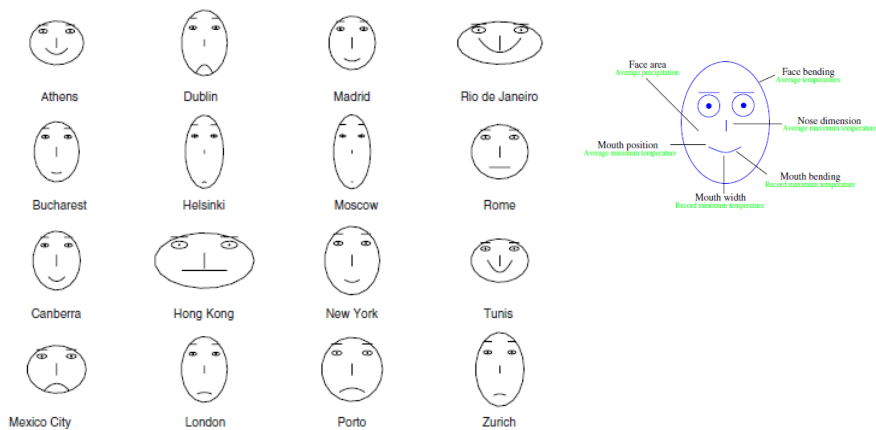
School of Information Systems



## High-dimensional categorical data

Project Name	Address	Type of Area	Property Type	Tenure	Completion Date	Type of Sale	Address Indicator
POH HENG COURT	1B Woo Mon Chew Road #02-01	Strata	Apartment	Freehold	1986	Resale	HDB
CHARLESTON	10 Shanghai Road #12-03	Strata	Apartment	Freehold	2001	Resale	Private
GREENRIDGE	21 Greenridge Crescent	Land	Terrace House	Freehold	1987	Resale	Private
BLUE HORIZON	27 West Coast Crescent #21-22	Strata	Condominium	99 Yrs From 14/08/2000	2005	Resale	HDB
PINEWOOD GARDENS	21 Balmoral Park #01-05	Strata	Condominium	Freehold	1990	Resale	Private
WOODSVALE	3 Woodlands Drive 72 #13-06	Strata	Executive Condominium	99 Yrs From 22/09/1997	2000	Resale	HDB
BAYVILLE CONDOMINIUM	42 South Buona Vista Road #04-12	Strata	Condominium	Freehold	1996	Resale	HDB
HILLVIEW GREEN	69 Hume Avenue #04-05	Strata	Condominium	999 Yrs From 19/05/1883	1998	Resale	Private
KATONG APARTMENTS	28C Mangs Road	Strata	Apartment	Freehold	1995	Resale	Private
SPOTTISWOODE RESIDENCES	48 Spottiswoode Park Road #26-12	Strata	Condominium	Freehold	Uncompleted	New Sale	HDB
THE CASCADIA	939 Bukit Timah Road #07-40	Strata	Condominium	Freehold	2010	New Sale	HDB
CHANGI RISE CONDOMINIUM	16 Simei Rise #07-40	Strata	Condominium	99 Yrs From 29/11/2000	2004	Resale	Private
SPOTTISWOODE RESIDENCES	48 Spottiswoode Park Road #26-05	Strata	Condominium	Freehold	Uncompleted	New Sale	HDB
THE TENNERY	5 Woodlands Road #06-20	Strata	Apartment	99 Yrs From 24/05/2010	Uncompleted	New Sale	HDB
THE LAKESHORE	29 Jurong West Street 41 #07-08	Strata	Condominium	99 Yrs From 22/11/2002	2008	Resale	HDB
WINDERMERE	28 Choa Chu Kang Street 64 #15-03	Strata	Executive Condominium	99 Yrs From 03/03/1997	1999	Resale	Private
THE CASCADIA	943 Bukit Timah Road #04-47	Strata	Condominium	Freehold	2010	New Sale	HDB
WATTEN HILL	52 Watten View	Strata	Condominium	Freehold	1979	Resale	Private
AMANINDA	328A Thomson Road #02-10	Strata	Condominium	Freehold	2005	Resale	Private
N.A.	104A Owen Road	Strata	Apartment	Freehold	1985	Resale	Private

## Chernoff Faces And Glyphs



Source: The Use of Faces to Represent Points in k-Dimensional Space Graphically (<http://www.jstor.org/stable/2284077>) and A Critique of Chernoff Faces (<http://eagereyes.org/criticism/chernoff-faces>)

## The Titanic Disaster - the tragedy



Source:  
[http://en.wikipedia.org/wiki/Sinking\\_of\\_the\\_RMS\\_Titanic](http://en.wikipedia.org/wiki/Sinking_of_the_RMS_Titanic)

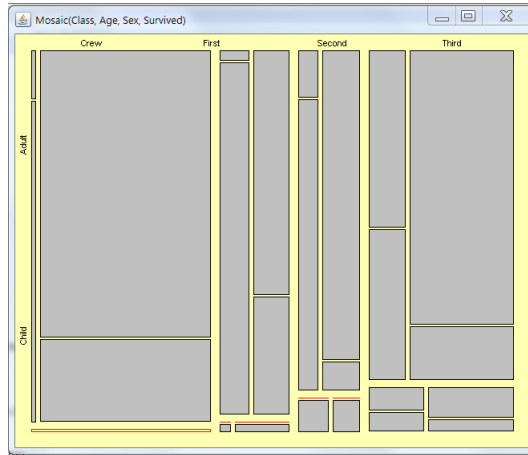
## The Titanic Disaster - the data

- The general rule
  - Women and children first

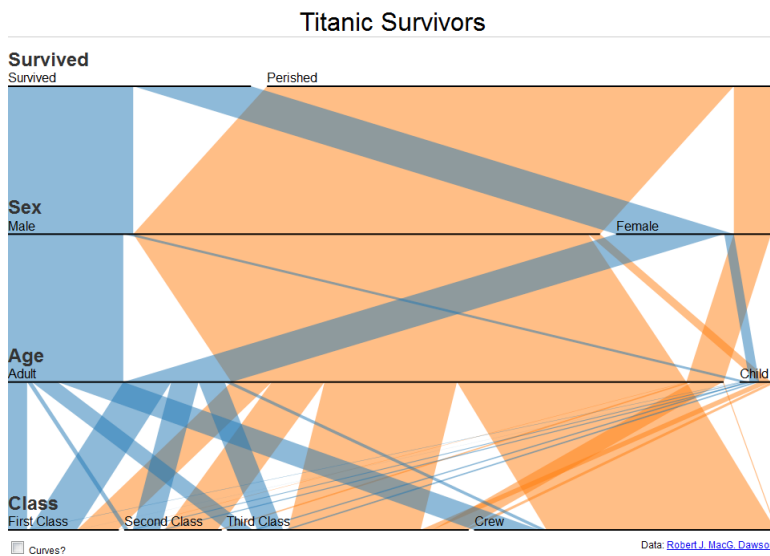
	Class	Age	Sex	Survived
1	First	Adult	Male	Yes
2	First	Adult	Male	Yes
3	First	Adult	Male	Yes
4	First	Adult	Male	Yes
5	First	Adult	Male	Yes
6	First	Adult	Male	Yes
7	First	Adult	Male	Yes
8	First	Adult	Male	Yes
9	First	Adult	Male	Yes
10	First	Adult	Male	Yes
11	First	Adult	Male	Yes
12	First	Adult	Male	Yes
13	First	Adult	Male	Yes
14	First	Adult	Male	Yes
15	First	Adult	Male	Yes
16	First	Adult	Male	Yes
17	First	Adult	Male	Yes
18	First	Adult	Male	Yes
19	First	Adult	Male	Yes
20	First	Adult	Male	Yes
21	First	Adult	Male	Yes
22	First	Adult	Male	Yes
23	First	Adult	Male	Yes
24	First	Adult	Male	Yes
25	First	Adult	Male	Yes

## Visualising high-dimensional categorical data

- Mosaic plot created using Mondrain

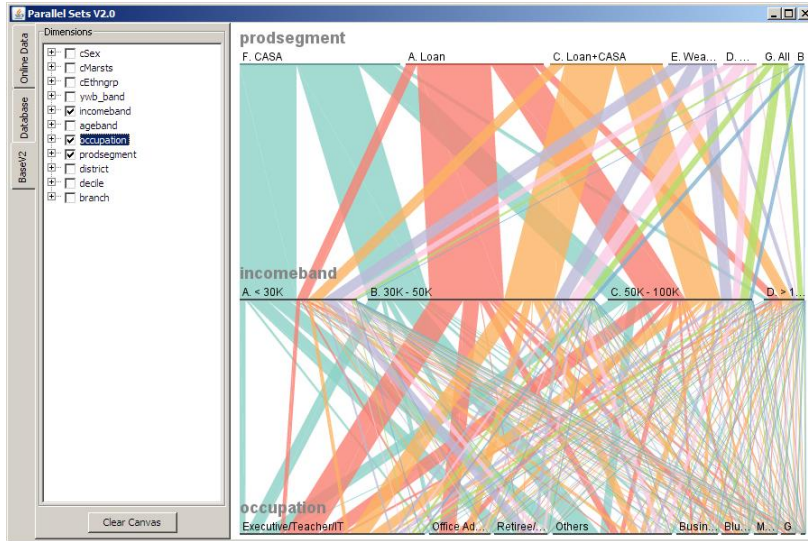


## Parallel Sets – d3.js





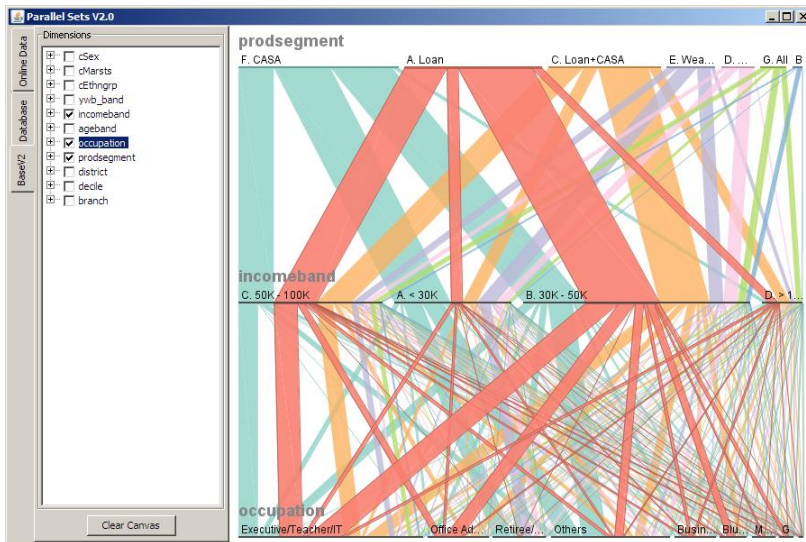
### Visualising High-dimensional Categorical Data with Parallel Sets



School of Information Systems



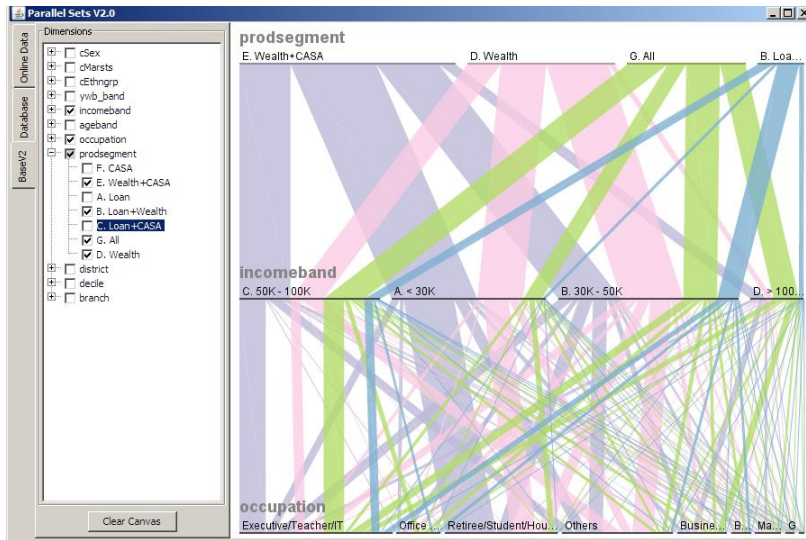
### Exploring High-dimensional Categorical Data with Highlighting Technique



School of Information Systems



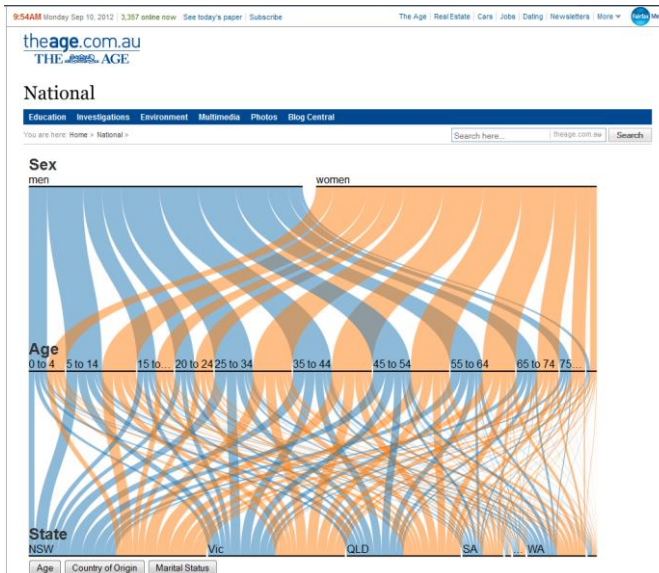
## Exploring High-dimensional Categorical Data with Filtering Technique



School of Information Systems



## Data Visualisation Application – Parallel Sets



School of Information Systems



## References

---

- Chernoff face ([https://en.wikipedia.org/wiki/Chernoff\\_face](https://en.wikipedia.org/wiki/Chernoff_face))
- Chernoff face in d3.js (<https://github.com/gnarmis/chernoff-faces>) and (<http://bl.ocks.org/larskotthoff/2011590>)
- Radar Chart ([http://en.wikipedia.org/wiki/Radar\\_chart](http://en.wikipedia.org/wiki/Radar_chart))
- Ternary plot ([https://en.wikipedia.org/wiki/Ternary\\_plot](https://en.wikipedia.org/wiki/Ternary_plot)), (<http://csmres.jmu.edu/geollab/Fichter/SedRx/readternary.html>) and (<http://wvaughan.org/ternaryplots.html>)
- ggtern 1.0.3.1 on CRAN
- Ternary Plot in D3.js (<http://bl.ocks.org/tomgp/7674234>), (<https://gist.github.com/widged/5780720>), (<http://bl.ocks.org/tomgp/7766353>)