

Two Examples of Correspondence Analysis

Hervé Abdi

`www.utdallas.edu~herve` `herve@utdallas.edu`

Refresher! and Menu

MCA ...

Home Page

Title Page

◀◀

▶▶

◀

▶

Page 1 of 20

Go Back

Full Screen

Close

Quit

1. Refresher! and Menu

- Tow examples of simple CA

Refresher! and Menu

MCA ...

Home Page

Title Page

◀◀

▶▶

◀

▶

Page 2 of 20

Go Back

Full Screen

Close

Quit

1. Refresher! and Menu

- Tow examples of simple CA

Refresher! and Menu

MCA ...

Home Page

Title Page

◀◀

▶▶

◀

▶

Page 2 of 20

Go Back

Full Screen

Close

Quit

2. MCA ...

Science Diplomas in the US from 1966 to 2000

Refresher! and Menu

MCA ...

Home Page

Title Page



Page 3 of 20

Go Back

Full Screen

Close

Quit

Table 1: Correspondence Analysis. All Diplomas in the US (1966–2000).
(Bachelor, Master, Ph.D., Men and Women) The Data

	Engineer	Physics	Earth	Math	Biology	Psychology	Social	NonSc
1966	51832	22323	2875	25958	38380	20528	75036	445797
1967	52725	23337	3322	28093	42167	23628	84657	482735
1968	55776	24732	3781	31136	46961	28625	98991	546948
1969	60106	26988	4544	36068	52894	37962	115479	623301
1970	63813	26862	4973	37441	56231	39636	127087	680978
1971	65129	27012	5322	35333	57349	44737	134767	739571
1972	65984	25331	6061	35717	60945	50993	140819	794075
1973	66698	24628	6468	35907	67704	56436	144492	826048
1974	61614	23968	7117	24897	75076	61370	143041	858499
1975	57966	23375	7005	31169	79753	61291	134957	862714
1976	57673	23238	7268	29218	83588	61105	130883	887420
1977	60017	23289	8001	28189	84915	59104	125678	888982
1978	68756	23467	8456	27305	83585	56306	124597	885420
1979	71242	23581	8501	27750	81295	54134	118754	879394
1980	77232	23399	8576	30163	78288	53472	116429	883807
1981	82696	23439	9153	34153	73814	52761	113399	885616
1982	87663	23448	9730	40745	71578	52647	114155	891868
1983	94337	22296	9881	46508	69303	52711	109004	898952
1984	99211	18035	10515	55899	65930	51705	106128	890765
1985	101710	22809	10405	65512	64423	51836	105161	887601
1986	101292	22553	9847	71108	63268	52426	107166	895253
1987	100207	22276	7342	69440	61163	54533	109064	902411
1988	97067	21313	6169	64741	59607	56376	111795	922555
1989	95233	21285	5723	60577	58256	60814	119090	954570
1990	93594	20754	5110	57293	59480	66607	123912	993414
1991	91370	21068	5037	54981	61895	71940	133670	938022
1992	92325	21863	5414	55122	67918	77141	141401	1082076
1993	95987	21813	5662	55537	73731	81073	143846	1112176
1994	97508	22328	6104	55549	79615	84714	143788	1122270
1995	97967	22960	6739	55293	86919	89159	139822	1110577
1996	97136	23591	6730	64203	95514	90359	139765	1123887
1997	94290	23149	6773	54212	99750	91901	138914	1049739
1998	92940	23046	6558	57136	102173	91264	138575	1162226
2000	90574	21496	6144	68853	100123	91981	141055	1230449

Refresher! and Menu

MCA ...

Home Page

Title Page

◀

▶

◀

▶

Page 4 of 20

Go Back

Full Screen

Close

Quit

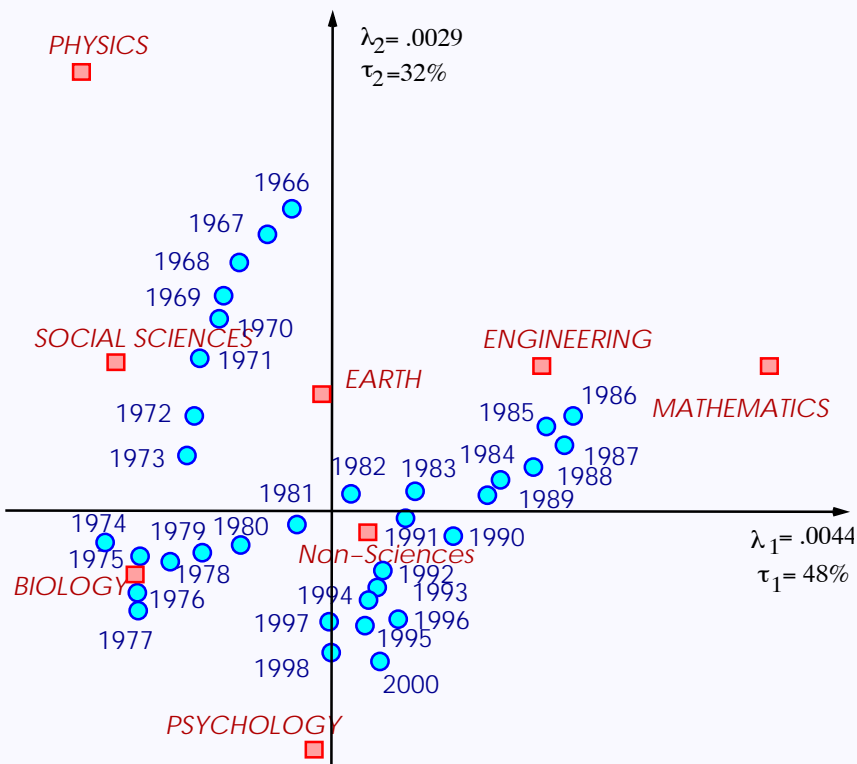


Figure 1: Diplomas in the US: Correspondence Analysis.

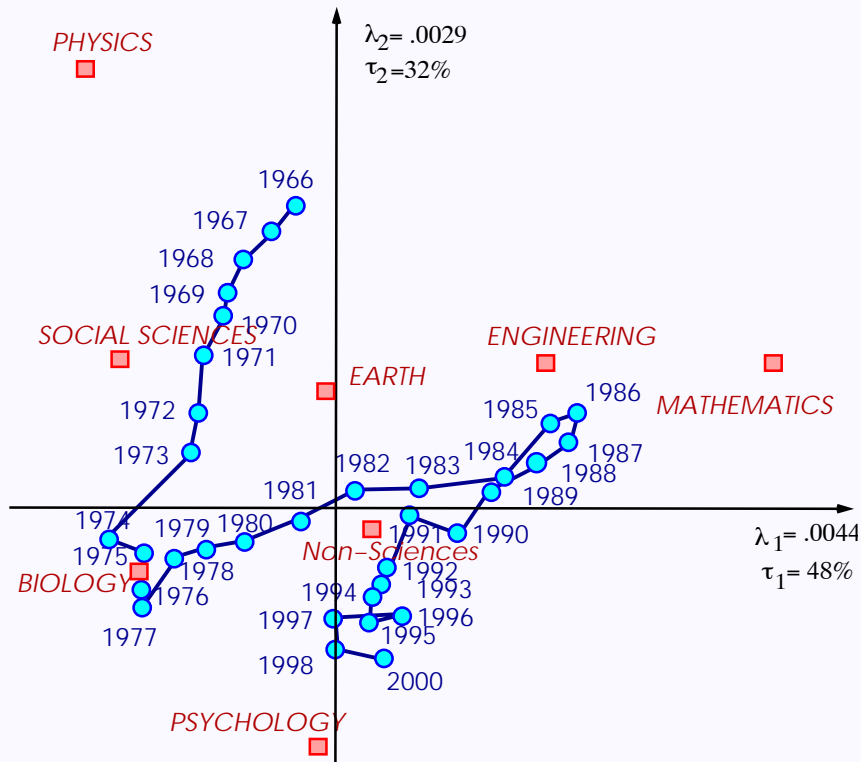


Figure 2: Diplomas in the US: Correspondence Analysis (with frills).

Refresher! and Menu

MCA ...

Home Page

Title Page

◀

▶

◀

▶

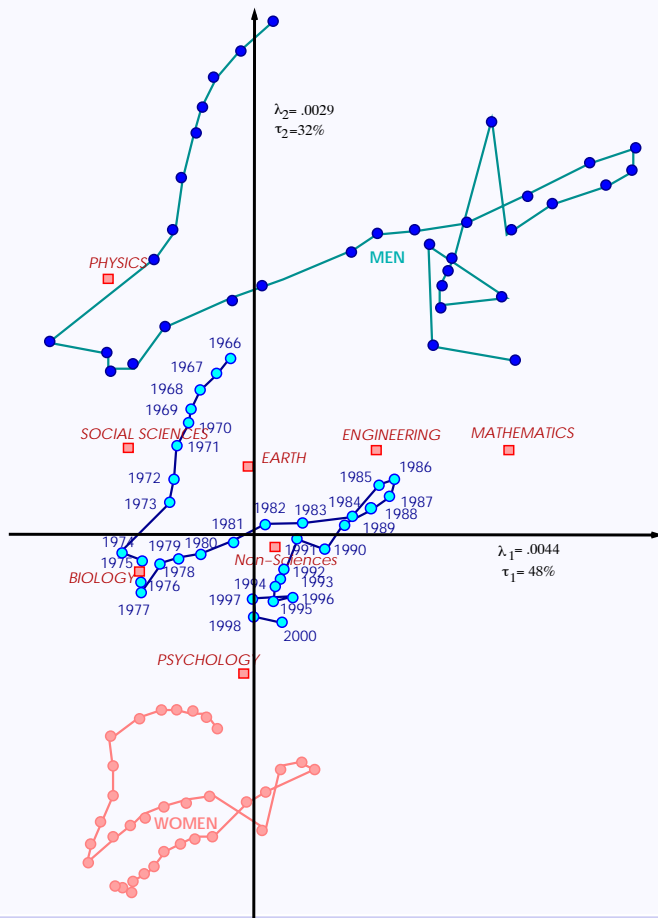
Page 6 of 20

Go Back

Full Screen

Close

Quit



Refresher! and Menu

MCA ...

Home Page

Title Page

◀◀

▶▶

◀

▶

Page 7 of 20

Go Back

Full Screen

Close

Quit

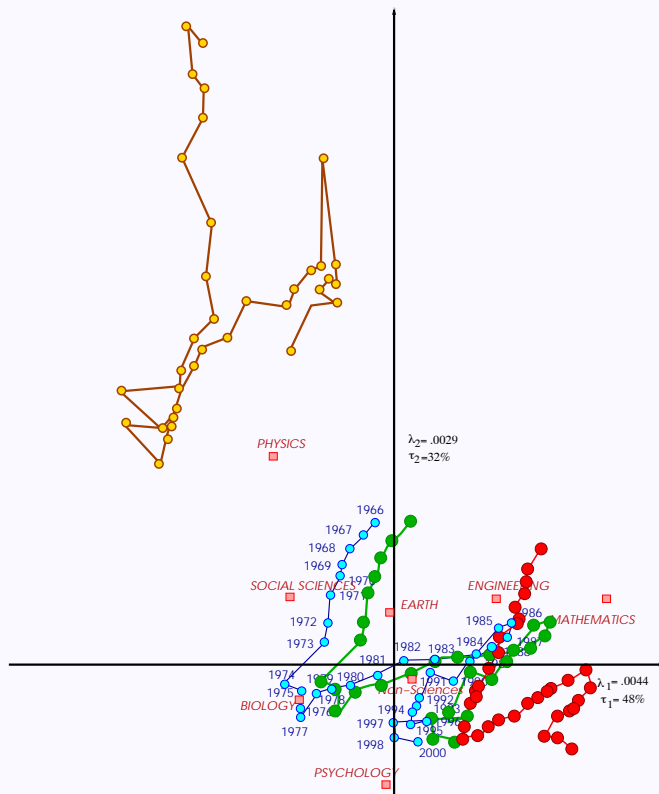


Figure 4: Diplomas in the US: Correspondence Analysis. By Diploma (Green: Bachelor. Red: Master. Gold: Ph.D.).

Home Page

Title Page

◀◀

▶▶

◀

▶

Page 8 of 20

Go Back

Full Screen

Close

Quit

Table 2: Correspondence Analysis. Diploma Global Analysis. J -set (Disciplines). The Projections, Contributions, Squared distance to barycenter, and Cosines

	F_1	F_2	ctr_1	ctr_2	$m \times d^2$	\cos_1^2	\cos_2^2	Qual
Engineer	0.11	0.07	63	26	.0208	18	12	30
Physics	-0.12	0.22	21	74	.0658	6	30	35
Earth	-0.00	0.06	0	4	.0770	0	1	1
Math	0.23	0.07	82	8	.0621	41	6	47
Biology	-0.10	-0.03	53	6	.0177	12	2	14
Psychology	-0.00	-0.12	0	72	.0202	0	23	23
Social	-0.10	0.08	55	32	.0196	23	21	44
NonSc	0.00	-0.02	3	70	.0004	0	7	7

Home Page

Title Page

◀

▶

◀

▶

Page 9 of 20

Go Back

Full Screen

Close

Quit

Table 3: Correspondence Analysis. Diploma Global Analysis. I -set (Disciplines). The Projections, Contributions, Squared distance to barycenter, and Cosines

	F_1	F_2	ctr_1	ctr_2	$10^3 \times$ $m \times d^2$	\cos_1^2	\cos_2^2	Qual
1966	-0.02	0.15	2	87	.3968	0	12	12
1967	-0.03	0.14	5	88	.3765	0	11	12
1968	-0.05	0.13	11	85	.3557	1	10	11
1969	-0.05	0.11	17	74	.3471	1	9	10
1970	-0.06	0.10	22	70	.3387	2	8	10
1971	-0.07	0.08	38	50	.2956	3	5	8
1972	-0.07	0.05	53	26	.2370	3	2	5
1973	-0.07	0.03	70	10	.2210	4	1	4
1974	-0.11	-0.01	94	1	.3978	9	0	9
1975	-0.10	-0.03	90	6	.2972	6	1	7
1976	-0.10	-0.04	82	13	.3432	6	2	8
1977	-0.09	-0.04	76	14	.3359	6	2	7
1978	-0.08	-0.03	72	8	.2682	4	1	5
1979	-0.07	-0.02	58	7	.2305	3	1	4
1980	-0.05	-0.02	38	5	.1633	1	0	2
1981	-0.02	-0.01	10	1	.1153	0	0	0
1982	0.01	0.01	1	2	.0881	0	0	0
1983	0.04	0.01	34	2	.1437	1	0	1
1984	0.09	0.01	69	2	.3073	5	0	5
1985	0.11	0.04	76	11	.4583	8	2	10
1986	0.12	0.05	81	11	.5413	10	2	12
1987	0.12	0.04	89	9	.4513	9	1	11
1988	0.10	0.02	93	4	.3195	7	0	7
1989	0.08	0.01	85	1	.2235	4	0	4
1990	0.06	-0.01	64	2	.1882	3	0	3
1991	0.04	-0.00	34	0	.1361	1	0	1
1992	0.03	-0.03	16	23	.1446	1	1	2
1993	0.02	-0.04	14	38	.1324	0	2	2
1994	0.02	-0.04	11	59	.1165	0	2	3
1995	0.01	-0.05	6	75	.1301	0	3	4
1996	0.02	-0.05	13	55	.1690	0	3	4
1997	-0.00	-0.05	0	45	.2342	0	4	4
1998	0.00	-0.07	0	85	.2264	0	7	7
2000	0.03	-0.08	8	69	.3213	1	8	8

Refresher! and Menu

MCA ...

Home Page

Title Page

◀

▶

◀

▶

Page 10 of 20

Go Back

Full Screen

Close

Quit

The Causes of Death in the USA in 2001.

Refresher! and Menu

MCA ...

Home Page

Title Page



Page 11 of 20

Go Back

Full Screen

Close

Quit

The Causes of Death in the USA in 2001.

Table 4: The Causes of Death in the USA (2001).

	-1	1-4	5-14	15-24	25-34	35-44	45-54	55-64	65-74	75-84	85+
Septicemia	312	108	69	133	281	829	1977	3111	5998	10351	9069
Hepatitis	1		3	20	74	797	2060	1014	836	613	166
hiv	10	13	38	225	2101	5867	4120	1317	384	88	10
Cancer	66	420	1008	1704	3994	16569	49562	90223	147018	165445	77751
tumor	51	58	105	102	151	366	719	1253	2574	4654	3643
Diabetes	2	6	39	151	595	1958	5343	9570	16731	22805	14171
Parkinson-											
Alzheimer	0	0	0	0	4	18	118	830	5569	26667	37190
Heart	600	285	365	1280	4038	16686	44450	76224	148409	291995	337930
Pneumonia	299	112	92	181	339	983	1801	2704	6650	18677	30191
Lower-											
Respiratory	41	43	104	171	290	969	3324	11166	30751	47762	28391
Liver	3		2	26	387	3336	7259	5750	5486	3801	979
Kidney	133	16	18	92	225	768	1822	3284	7356	13097	12668
Unknown	3491	272	181	949	1530	3126	3504	2292	2774	5003	9089
Other											
Diseases	1131	646	860	1799	3072	8084	14203	16048	25803	49711	65032
Accident-											
Car	151	678	1823	11139	7454	7827	6251	3861	3282	3507	1272
Accident-											
NonCar	144	626	1686	10725	6937	7083	5564	3465	2990	3312	1233
Other											
Accidents	825	1036	1013	3272	4385	8118	7093	3797	4553	9181	10899
Suicide	0	0	279	3971	5070	6635	5942	3317	2432	2192	769
Homicide	332	415	326	5297	5204	4268	2467	1018	532	312	105

Refresher! and Menu

MCA ...

Home Page

Title Page

◀

▶

◀

▶

Page 11 of 20

Go Back

Full Screen

Close

Quit

First have a look at the data (log transformed):

Refresher! and Menu

MCA ...

Home Page

Title Page



Page 12 of 20

Go Back

Full Screen

Close

Quit

First have a look at the data (log transformed):

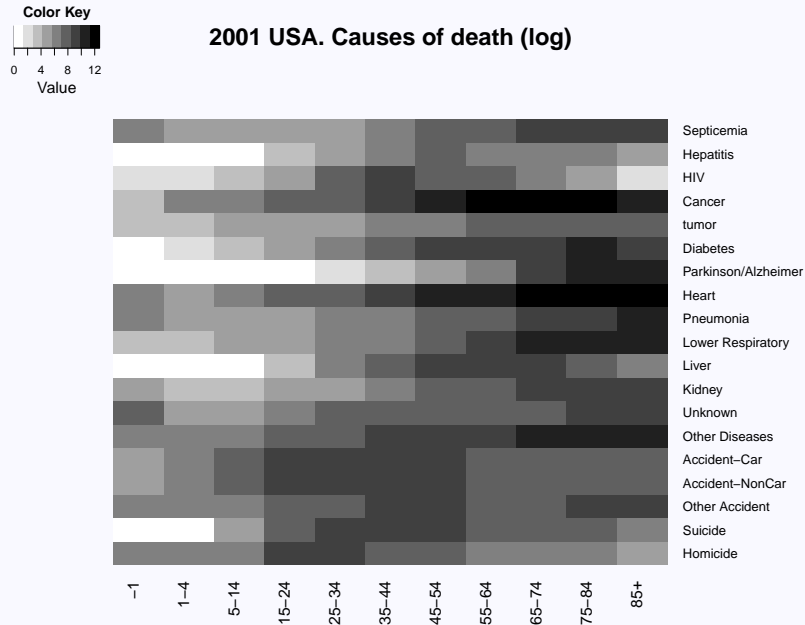


Figure 5: Causes of Death in the US: A picture of the data (log transformed).

Refresher! and Menu

MCA ...

Home Page

Title Page

◀◀

▶▶

◀

▶

Page 12 of 20

Go Back

Full Screen

Close

Quit

Correspondence analysis: The Scree.

Refresher! and Menu

MCA ...

Home Page

Title Page



Page 13 of 20

Go Back

Full Screen

Close

Quit

Correspondence analysis: The Scree.

Refresher! and Menu

MCA ...

Home Page

Title Page



Page 13 of 20

Go Back

Full Screen

Close

Quit

Explained Variance per Dimension

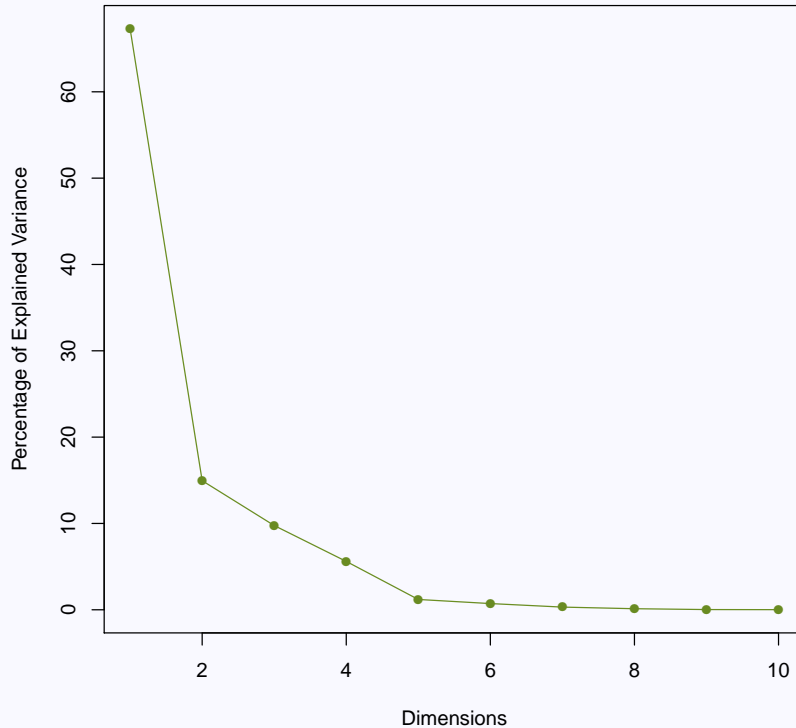


Figure 6: Causes of Death in the US: The Scree (How many dimensions?).

Refresher! and Menu

MCA ...

Home Page

Title Page

◀◀

▶▶

◀

▶

Page 14 of 20

Go Back

Full Screen

Close

Quit

A (symmetric) plot

Refresher! and Menu

MCA ...

Home Page

Title Page



Page 15 of 20

Go Back

Full Screen

Close

Quit

A (symmetric) plot

Refresher! and Menu

MCA ...

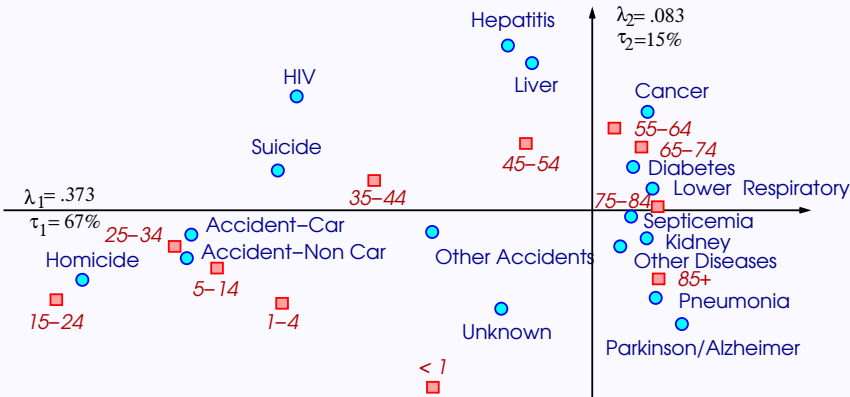


Figure 7: Causes of Death in the US: Correspondence Analysis. Symmetric plot. $\lambda_1 = .3730$, $\lambda_2 = .0829$.

Home Page

Title Page

◀

▶

◀

▶

Page 15 of 20

Go Back

Full Screen

Close

Quit

A (symmetric) plot

Refresher! and Menu

MCA ...

Home Page

Title Page



Page 16 of 20

Go Back

Full Screen

Close

Quit

A (symmetric) plot

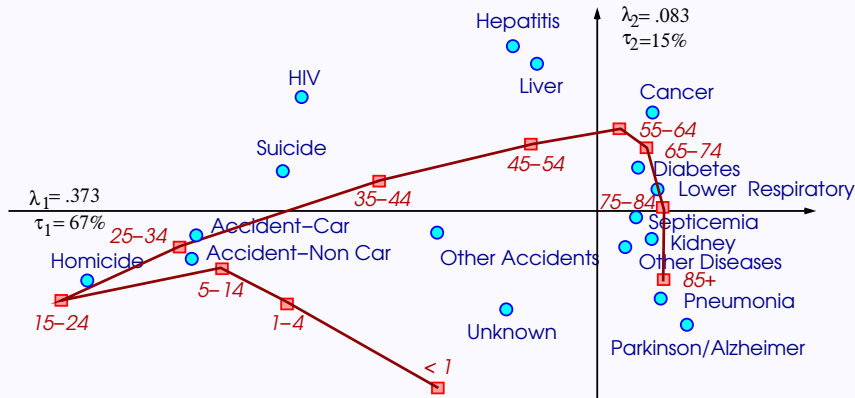


Figure 8: Causes of Death in the US: Correspondence Analysis. Symmetric plot with frills. There are three periods in life (or death!). $\lambda_1 = .3730$, $\lambda_2 = .0829$.

An asymmetric plot from R

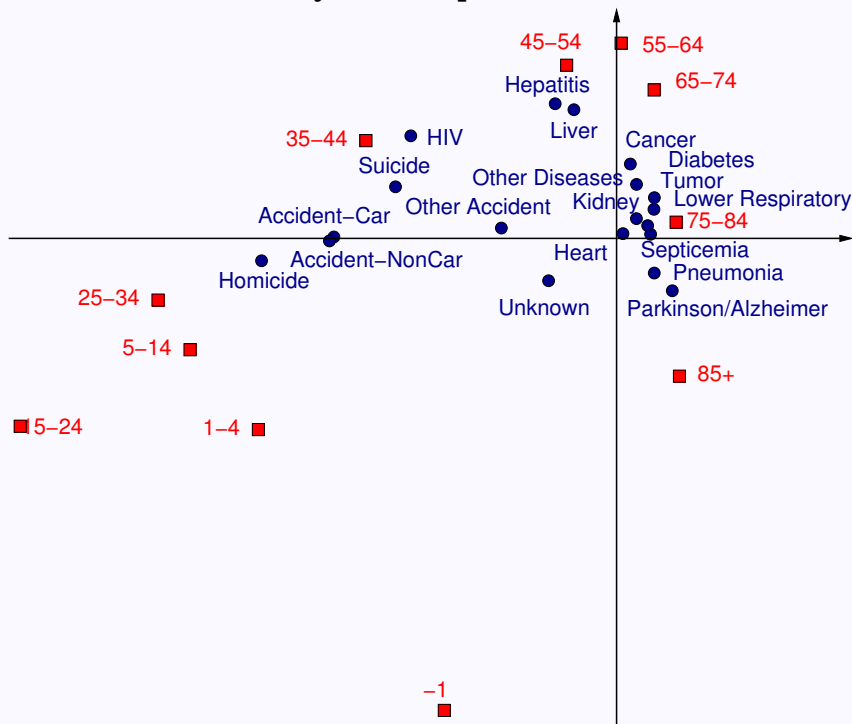


Figure 9: Causes of Death in the US: Correspondence Analysis. Asymmetric plot (From R). There are three periods in life (or death!). $\lambda_1 = .3730$, $\lambda_2 = .0829$.

Refresher! and Menu

MCA ...

Home Page

Title Page

◀

▶

◀

▶

Page 17 of 20

Go Back

Full Screen

Close

Quit

An asymmetric plot (with frills)

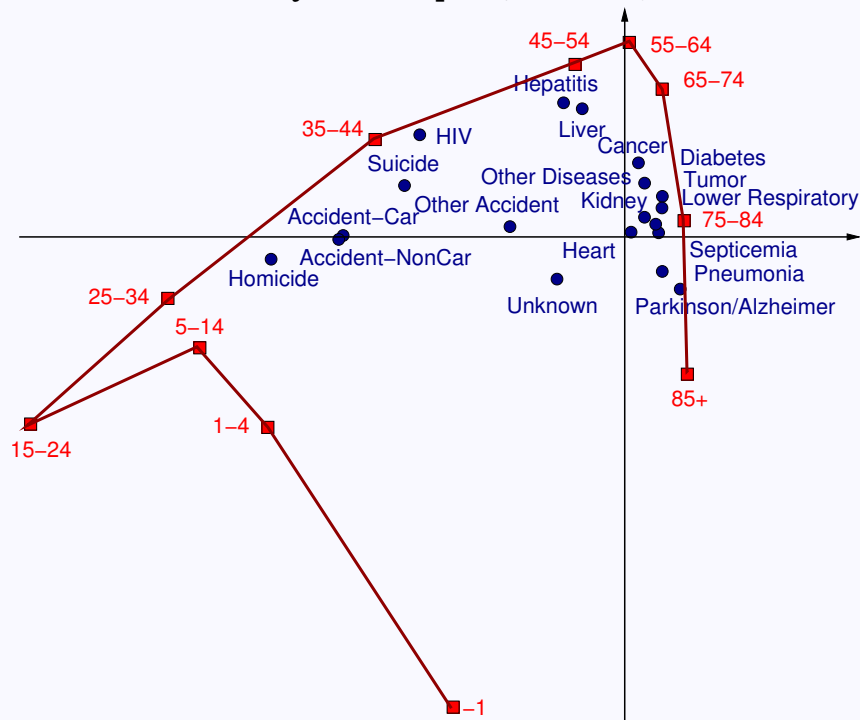


Figure 10: Causes of Death in the US: Correspondence Analysis. Asymmetric plot with frills (From R). There are three periods in life (or death!). $\lambda_1 = .3730$, $\lambda_2 = .0829$.

Refresher! and Menu

MCA ...

Home Page

Title Page

◀

▶

◀

▶

Page 18 of 20

Go Back

Full Screen

Close

Quit

	F_1	F_2	ctr ₁	ctr ₂	$m \times d^2$	\cos_1^2	\cos_2^2	Qual
-1	-0.82	-1.10	4	8	0.05	1	5	5
1-4	-1.69	-0.48	63	5	0.01	2	1	2
5-14	-2.01	-0.30	77	2	0.02	4	0	4
15-24	-2.88	-0.47	90	2	0.16	39	5	44
25-34	-2.16	-0.19	96	1	0.10	25	1	26
35-44	-1.18	0.16	79	2	0.07	15	1	16
45-54	-0.40	0.35	41	31	0.03	3	11	14
55-64	0.02	0.38	0	92	0.02	0	18	18
65-74	0.17	0.28	23	60	0.02	1	17	18
75-84	0.24	-0.02	86	0	0.02	5	0	5
85-	0.29	-0.36	38	58	0.06	6	42	49

Table 5: Correspondence Analysis. Causes of Death in the US. Global Analysis. J -set (Disciplines). The Projections, Contributions, Squared Distances to Barycenter, and (squared) Cosines.

[Home Page](#)
[Title Page](#)
[◀](#)
[▶](#)
[◀](#)
[▶](#)
[Page 19 of 20](#)
[Go Back](#)
[Full Screen](#)
[Close](#)
[Quit](#)

	F_1	F_2	ctr_1	ctr_2	$m \times d^2$	\cos_1^2	\cos_2^2	Qual
Septicemia	0.15	-0.03	52	2	0	0	0	0
Hepatitis	-0.48	0.86	12	38	0	0	2	2
HIV	-1.59	0.61	45	7	0.03	4	3	7
Cancer	0.10	0.39	6	88	0.04	1	44	44
Tumor	0.12	-0.02	35	1	0	0	0	0
Diabetes	0.15	0.23	26	67	0	0	2	2
Parkinson/Alzheimer	0.42	-0.59	32	62	0.02	1	12	14
Heart	0.26	-0.15	71	24	0.04	7	11	18
Pneumonia	0.28	-0.45	27	69	0.01	1	6	7
Lower Respiratory	0.29	0.08	51	4	0.01	1	0	2
Liver	-0.34	0.81	10	57	0.01	0	9	9
Kidney	0.23	-0.08	84	11	0	0	0	0
Unknown	-0.53	-0.51	8	7	0.05	1	4	5
Other Diseases	0.04	-0.15	4	49	0	0	2	2
Car Accidents	-2.19	-0.17	95	1	0.10	26	1	27
Non-car Accidents	-2.22	-0.20	94	1	0.10	25	1	26
Other Accidents	-0.89	-0.10	78	1	0.02	5	0	5
Suicide	-1.71	0.22	90	1	0.04	10	1	11
Homicide	-2.75	-0.36	95	2	0.07	17	1	19

Table 6: Correspondence Analysis. Cause of Death in the US Global Analysis. I -set (Causes). The Projections, Contributions, Squared distance to barycenter, and Cosines

Refresher! and Menu

MCA ...

Home Page

Title Page

◀

▶

◀

▶

Page 20 of 20

Go Back

Full Screen

Close

Quit