Comments for MEDB 5502, Week 05

Topics to be covered

- What you will learn
 - Principal components analysis
 - Applications of principal components
 - Factor analysis
 - Criticisms of principal components analysis and factor analysis

What is a linear combination?

- Multiply each variable by a constant and add
- Examples

$$X_1 + X_2 + X_3 + X_4$$

$$-\frac{1}{2} \times X_1 + 0X_2 + 0X_3 + \frac{1}{2}X_4$$

$$-\frac{1}{3}X_1 + \frac{1}{3}X_2 + \frac{1}{3}X_3 - 1 \times X_4$$

$$-9X_1+3X_2+5X_3-X_4$$

Speaker notes

- The simplest linear combination is just a sum. Multiply each variable by one and add the values together.
- Almost as simple is an average. With four variables, multiply each variable by one fourth and add the values together.
- A linear combination could effectively exclude one or more of the variables just by multiplying them by zero. The third linear combination is an average of just the first and last variables.
- You can put negative numbers in the linear combination. The fourth linear combination computes the average of the first three variables and subtracts the last variable.
- The fifth linear combination is a weighted average. The first two variables are weighted four times as much as the last two variables.
- A linear combination does not necessarily have to involve averages. Taking 9 times the first variable plus 3 times the second variable plus 5 times the third variable minus the fourth variable.

Not a linear combination?

- $X_1 \times X_2 + X_3 + X_4$
- $X_1^4 + X_2 + X_3 + X_4$
- $log(X_1) + X_2 + X_3 + X_4$

Speaker notes

Multiply two or more of the variables together and it is no longer a linear combination. Raise one of the variables to a power and it is no longer a linear combination. Taking a function of one of the variables and it is no longer a linear combination.

Practical examples

- Apgar score is a linear combination
 - Rate appearance, pulse, grimace response, activity, and respiration
 - Add the values together
- Body mass index is not a linear combination
 - Weight divided by height squared

Speaker notes

Apgar score is a linear combination. Rate five features of a newborn baby on a scale of 0, 1, or 2. Then add the five values together to produce a linear combination that equals zero for a baby in extreme distress all the way to a happy baby with an apgar score of 10. Actually, the baby has to have a loud cry to get a 10, so perhaps this is more of a healthy baby than a happy baby.

Body mass index is not a linear combination. It involved division, which is just as much of a no-no as multiplication. It also involve a power of two which would also make it not a linear combination.

Philosophy behind principal components, 1 of 4

- Reduce complexity by modeling inter-relationships
- Inter-relationships represented by linear combinations
 - There is no dependent or outcome variable in principal components analysis

Philosophy behind principal components, 2 of 4

- First principal component
 - Linear combination that accounts most variation
 - This linear combination is the first eigenvector
 - The amount of variation accounted for is the first eigenvalue

Philosophy behind principal components, 3 of 4

- Need to resolve an ambiguity
 - $lacksquare 3X_1 + 5X_2 4X_3 + 7X_4 1X_5$ versus $6X_1 + 10X_2 8X_3 + 14X_4 2X_5$
 - Solution: require sum of squared coefficients to equal 1
 - Note: $3^2 + 5^2 + (-4)^2 + 7^2 + (-1)^2 = 100$
 - $lacksquare Use rac{3}{10} X_1 + rac{5}{10} X_2 rac{4}{10} X_3 + rac{7}{10} X_4 rac{1}{10} X_5$

Philosophy behind principal components, 4 of 4

- Second principal component
 - Linear combination that accounts second most variation
 - Must be uncorrelated with first principal component
 - This linear combination is the second eigenvector
 - The amount of variation accounted for is the second eigenvalue
- Third principal component defined similarly

Covariance matrix or correlation matrix

- Correlation matrix equivalent to standardizing
 - Absolute requirement if differing units
- Covariance matrix de-emphasizes low variance variables

How many components?

- Percentage of variation accounted for
 - Scree plot
- Eigenvalues > 1
- Researcher preference/convenience

Communality

- Amount of shared variation
 - Always between 0 and 1
 - Similar interpretation to R-squared
 - "One of these things is not like the others"

Factor score matrix

- Linear combination coefficients
- Needed if you score by hand
- No obvious interpretation
- First component is often only positive values

Component matrix

- Interpret as correlation matrix
 - Rows are individual variables
 - Columns are principal components

Health-related quality of life, 1 of 6

```
data_dictionary: Health-related Quality of Life of Persons Living with
Tuberculosis.sav

source: |
   Tornu, Eric ; Quarcoopome, Louisa (2022), "Health-related Quality of Life
of Persons Living with Tuberculosis", Mendeley Data, V2, doi:
10.17632/jg4xp7883w.2

description: |
   From the original source: This data is on a study which accessed the health-
related quality of life of persons living with Tuberculosis.
```

Speaker notes Here are a few pieces of the data dictionary for a file I will be using. I won't show the entire data dictionary, as it is quite long.

Health-related quality of life, 2 of 6

```
Howwouldyourateyourqualityoflife 1:
 values:
    '1': Poor
    '2': Neither poor nor good
    '3': Good
    '4': Very Good
HowSatisareyouwithyourhealth 2:
 values:
    '1': Very dissatisfied
    '2': Dissatisfied
    '3': Neither satisfied nor satisfied
    '4': Satisfied
    '5': Very Satisfied
Physicalpainprevents 3:
 values:
    '1': An extreme amount
    '2': Very much
    '3': A moderate amount
    '4': A little
    '5': Not at all
```

Health-related quality of life, 3 of 6

```
Howmuchmedicaltreatment 4:
  values:
    '1': An extreme amount
    '2': Very much
    '3': A moderate amount
    '4': A little
    '5': Not at all
Howmuchdoyouenjoylife 5:
  values:
    '1': Not at all
    '2': A little
    '3': A moderate amount
    '4': Very much
    '5': Extremely
Towhatextentdoyoufeelyourlifetobemeaningful 6:
  values:
    '1': Not at all
    '2': A little
    '3': A moderate amount
    '4': Very much
    '5': Extremely
```

Health-related quality of life, 4 of 6

```
Howwellareyouabletoconcentrate 7:
  values:
    '1': Not at all
    '2': A little
    '3': A moderate amount
    '4': Very much
    '5': Extremely
Howsafedoyoufeelinyourdailylife 8:
  values:
    '1': Not at all
    '2': A little
    '3': A moderate amount
    '4': Very much
    '5': Extremely
Howhealthyisyourphysicalenvironment 9:
  values:
    '1': Not at all
    '2': A little
    '3': A moderate amount
    '4': Very much
    '5': Extremely
```

Health-related quality of life, 5 of 6

```
Doyouhaveenoughenergyforeverydaylife 10:
  values:
    '1': Not at all
    '2': A little
    '3': A moderate amount
    '4': Very much
    '5': Extremely
Areyouabletoacceptyourbodilyappearance 11:
  values:
    '1': Not at all
    '2': A little
    '3': A moderate amount
    '4': Very much
    '5': Extremely
Haveyouenoughmoneytomeetyourneeds 12:
  values:
    '1': Not at all
    '2': A little
    '3': A moderate amount
    '4': Very much
    '5': Extremely
```

Health-related quality of life, 6 of 6

```
AvailableInformation 13:
  values:
    '1': Not at all
    '2': A little
    '3': A moderate amount
    '4': Very much
    '5': Extremely
Opportunityforleisureactivities 14:
  values:
    '1': Not at all
    '2': A little
    '3': A moderate amount
    '4': Very much
    '5': Extremely
Howwellareyouabletogetaround 15:
  values:
    '1': Very poor
    '2': Poor
    '3': Neither poor nor good
    '4': Good
    '5': Very Good
```

Correlation matrix, 1 of 3

			64	endature.																					
		CONTRACTOR OF THE PARTY OF THE	March of B	New York Street	To all all and a second of the	reference des	MONTHS AND	reductivatilly to		Annay day		The controller Topics for Microglan Transporters Transporters Transporters	To solve under the property of	Tolker trade (etc.) (the state of the part (second)	Profitable op-	Total Miles	recording on the old start reports to	Arra (Mill) pro provide provide	THE TARK OR	Name and Address of the Owner o	The late on particular to the late of the	The late or acceptance or acceptance or processors	Place Sales are any arthursor archaracte fraueth	market or partition	Indian Alex III. Into Takes Incipation Benings Auch
To all out of the design to the life.		- 1	.00		-		-	- 20	-	-	- 10		-346	-	100					-	107			- 10	- 10
Printed partyresetta pro-			1.00	1.00	140	1401	100	100		- 100	-	-	-00	- 100	-00	1.000	- 200	1.000	244		Total Control	.00	107	746	
Management and State and St. State of St. St	Parantinophia	300	318	360	39	340	100	- 39	-636	28	- 194	- 176	- 200	76	200	26	500	- 20	200	76	200	- 39	- 30	- 340	
	THE PERSON NAMED IN	- 190		- 10	- 10			1,000	10	1010	- 100	100	1,000	- 12	80		500	- 20		197	907	100	100	1,000	80
London Company		- 100	200		10	- 10	- 10	- 20	- 10	.796	- 10	- 10	100	100	100	- 100	- 10	-		- 10	- 10	- 10	- 100	- 10	
the hubble as yet	Propositionistics	307	- 10	-	.00		200	- 100	- 10	100	450	-	- 00		- 10	- 100	100		- 10	- 40	- 100			.00	
EU.	Sty. Drodest	1001	- 81	200	100	1000	100	100	-88	1,000	100	1,000	- 19	100	100	1,00	100	1,000	100	1,000	100	1,000	100	100	100
No entered in the second liver	Proposition	- 1	200	756	- 7	1600	596	-	- 10	- 10	400	- 46	- 600	100	579	- 100	860	- 00		- 41	-	200	100	.01	1871
Note the resemption	Sty. (Created)	100	90	100	200	1000	100	- 100	100	4,000	100	100	.716	100	100	100	100	1,000	100	100	100	100	100	100	100
toward stress areas	Proposition and Proposition in Contract of	107	- 10	- 70	- 0	-	- 11	- 10	10	- 100	- 71	- 00	- 2	76	- 5	- 0	- 1	- 5		- 61	100	- 81	- 10	- 60	1070
LORDON PROPERTY.	Tip (Product	1301	20	-81	100		100	-30		-389		- 101	- 10	- 100	- 100	- 10	100	100	100	100	100	100	100	100	-30
to a table to a substitute to	Prevention beauty		- 10	766		75%	- 1	40	- 10	- 171	- 10	- 46	- 01	- 65	200	- 10	100	- 0	30	- 10	344	- 40	107	46	print.
MALES .	Sty Draine	100	- 10	100	100	-301	200	-30	- 10	-301		1.00	- 07	100	100	100	100	100	100	100	100	100	1.00	100	100
to both more	Proposition in the last	- 00	- 6	- 4	- 40	- 61	- 67		- 10	- 10	- 64	- 10	- 65	- 10	176		- 10	-	- 1	- 4	- 29	-	- 6	-	100
Printed and company	Sty Drafted	31	100	-30	100	-301	100	26	- 10	-301		- 101	-00	- 100	100	- 100	-00	1.00	1,00	- 10	100	100	1,000	100	- 30
Strong Specimens	James Continue		101	100	.04	100	447	.00	- 1	- 100	- 69	-	100	76	- 100	- 100		- 100	_	- 40	100	300	100	- 10	360
PROPERTY OF	No. Disability	100	- 63	- 30	100	-301	-80	-20	200	- 100		- 101	- 60	- 100	100	100	100	1.00	100	100	100	100	100	100	100
Name and Association	Presentantian.		- 61	- 61	- 87	- 2	.03		- 15		- 1	- 7	147		- 54	- 20	340	- 81	- 6	- 6	100	- 2	-	- 40	
your debatement.	Stag Stradest	- 100	39	-00	-80	-301	- 100	-30	- 20	200		- 26	-00	- 30	-00	- 100	-00	- 100	-00	- 100	-00	- 100	-30	- 30	100
Service complete service	Personal annual		-39.	100	- 40	100	- 10	40	439	700	-	-		100	.00		.00	- 0				.00	100	.001	1007
Notice restr	Jing Strained	-	100	-00	100	- 400	100	- 100	- 1	- 100	200	- 10	- 10		- 100	1.00	- 40	1.00	- 100	1.00	100	1,000	100	- 100	100
promoted more	Personal annial con-	-	-63		40	- 100	- 10	80	-		- 100	-	.00	- 100	80		- 10	-	363	- 1	- 10	- 10	-	- 10	100
Promoter Measurement	NI STATE	901	- 1	181		100		- 10	- 1	- 100			- 10		- 10	- 1			- 10	100	100	100	100	- 100	- 100
The advantage of the Santage Santage	Participation of the last	-38	- 20	500	20	37	20	100	740	360	200	76		76	- 60	.76	100	760	100	200	190		100	- 10	360
No. As	The Street	-90		- 10	21	-001	607	-30	611	- 100	-	- 100			-60	- 100	-60	- 100	-00	- 100	100	100	+00	-	-000
COMMENT OF STREET	Name of Street Street	100	200	79	80	200	474	200	740	200		79	200	150	176	750	80	200	- 0	194	279	366	100	- 20	200
per processed	Tip (Frank)	-90	100	100	160	- 400	160	- 100	- 40	- 100	- 10	- 10	-6		100	- 1	100	- 100	- 10	- 100	100	1,000	100	1000	100
STATE STATE OF STREET	N. Proposition (Section)	38	200	- 2	570	200	- 20	200		200	- 10	76	- 10	760	- 100		- 2	760	- 2	360	100	200	100	- 60	360
H	Tip (Challed)	100	- 40	-00	- 40	-000	- 40	100	- 4	- 100		1.00	-01	- 6		1.00	100	- 100	1.00	- 100	100	- 100	1.00	100	-00
the fifth an authors	A	- 2	- 10	- 24	- 2	- 2	- 2	- 20	- 2	24	- 2	74	170	- 1	- 10		- 10	74	- 20	260		30	- 1	- 10	300
BURNING STATES	Sty Droines	190		-81	100	-30	100	-30	- 100	- 400	100	- 100	100	- 100	-50		100	1,000	1,00	1.00	100	100	1.00	1,000	1,000
the fall and another a	A. Carrier	- 2	- 10	- 2	- 20	- 2	- 10	20	- 2	- 34	- 10	- 24	- 10	240	- 20		-	- 20	100	- 20	100	- 20	- 10	- 2	100
reporters with	Tip Drames	100	.00	-30	100	-301	168	-30	100	- 100	-60	- 100	100	1,000	100	1,000		1.00	100	100	100	100	1.00	1,000	1000
the Street County	Towns Constitute	200	- 10		- 20	- 24	- 10		- 20	290	200 670	- 24	200	246	- 2		- 10		170	- 26	200	- 20	- 2	- 20	100
part of	Dy Drobet	130	- 11	-30	1000	1301	100	-30	100	1,000	100	1.00	100	1,000	100	1,000	100		100	100	- 100	1,000	100	100	1000
The Sales and Association of	Proposition in the last	- 34	200	- 20	39	295	200	28	200		170		200	296	- 2		- 10	360		200	200	36	200	30	390
personal analogophics	Tip, (Product	20	.00	1301	168	-30	100	-30	100	188	168	1,00	100	1,000	160	1,000	100	1,000		1,000	100	1,000	100	100	100
to design and the	Towns Contract	300	219	200	200	290	200	28	200	276	200	26	- 20	246	200	.76	- 2	200	200		200	360	- 20	30	36
10.00	Sq. Draine	301	160	-80	100	-80	-88	- 0	- 40	- 100	- 60	- 80	- 40	- 100	-88	1.00	- 60	1,000	1.00		- 100	1,000	- 10	100	100
to delice and other	Personal conductors	- 20	200	200	20	390	200		200	200	- 20	24	100	290	- 20		- 20	760	200	260	300	- 20	30	30	200
Section of the participation of the last o	The Street	100		140	160	1000	160	- 101	100	180	140	1,00	160	1,000	140	1,000	1,000	1,000	100	1.00		1,000	100	1,000	100
to delice and other	Townson Completion	- 30	- 29	- 20	20	200	200	246	200	200	700	76	- 20	246	200	260	- 20	700	170	200	100	-	200	200	200
STATE OF STREET	Sty. (Product		100	- 101	140	180	- 40	180	- 40	- 100	- 40	- 100	- 40	1,000	160	1,000	1.00	1.00	100	1.00	100		100	100	100
to the second or	N. Propose Constitute	- 24	20	21	30	24	200	24	200	76	500	76	200	74	200		- 27	- 26	100	36	100	200	- 10	300	200
SCHOOL SELECTION	Na Product	- 10	188	100	1488	100	- 40	180	- 40	188	100	1,000	- 40	1,000	100	1,000	100	1,000	100	1,000	100	1,000		1,000	1000
to de la constitución de la cons		- 51	- 26	20	21	- 24	100	24	200	- 24	- 10	76	200	74	40	76	- 10	200	100	50	80	30	100	- 10	300
THIS CO.	Title (Charles)		100	100	198	1000	198	1000	100	1,000		1,000	- 10	1,000	100	1,000	100	1,000	100	1,000	100	100	100		1001
Non-Mandage (See	N. Proposition (Application)	- 2	100		- 10	- 20	- 10	300	- 2	200	- 10	76	100	740	100	- 24	100	200	- 2	300	200	90	100	300	-
Name and Post Office of Street, or other Post Office of Street	Da China	82	.01	-00	-88	-90	-60	-80	-0	- 80	-00	- 10	-88	- 100	-98	- 100	-00	- 100	-00	- 100	-00	- 00	-00		
Married Street, and the		386	798		79	.00	700	(%)	700	196	100	(%)	-	(94)	-	(%)	- 10	(84)			-	760			200



Correlation matrix, 2 of 3

			Secretary (No.																				
	Townsome a production production the con-	medical residual residual is residual	The state of the s	To administrative Monte of the Park Total Control Con	Programme of the contract of t	THE PART OF THE PARTY.	Bran Tool Marry one of the di one of the di	Discourage emails emails emails emails		19		To attractions through from the secondarity for secondarity arrestment	2012/01/2015	Tray flat or to	1000	Accordingly for all county for regionally for agents	THE TANK THE PER	Now Tolks and story Williams posterior	Tel 1970	100 miles		Promision and promision and promision promision market	TOW THE ST.	
To select estant de province final	10	- 1					- 4					- 4		- 1			- 1	- 1	- 1		1			
State		100	- 1				- 48	- 1		- 4	- 10	- 4					- 1				- 4	- 0		
ery months of longerous for fundament or year (MA) (Ma)																								
THE PART OF STREET		- 1	1.0					- 1				- 1					- 7							
To select extent de province		-		14													-							
The facility becoming \$4.7.									,															
CALL STREET												- 1												
CONTRACTOR SOUTH			,	,		- 14																		
Total modified by the and a stranger of the st		- 0					- 10														- 1			
On you have interrupts amongs their								19				- 1	- 7											
Annual Santana	-		- 1						- 11			-		- 1										
took summer many transits		- 1								14		- 4	- 4				- 1				- 4			
The agreement assertion the agreement assertion thermalist flat yet found											- 11													
one doubt to the!												14					-							
Signaturilly for the latest of																								
Transactives provide the per-			,	,				,						,										
- Overlating rymunificacy of and								- 1						10							- 1			
CONTROL COLOR APPLICATION OF THE PERSON OF T				- 1	,			,		-					- 10									
the late error others.	- 1	-		- 1			- 1	- 1			-	- 1				- 4	- 1	-	- 1	- 4	- 1			
cope (b) for control with	- 1	-						- 1				- 1					- 11							
DUTCH HOWEVER DOWN ARTHUR HOWEVER BENEVICE	- 2								-		- 4	- 1		- 1			-	10						
the fall sector officer	- 1	- 1		- 4				- 1											- 1					
Total latin are on articles												-					-		-	- 10				
Sensor How Site service with the Continent year long	-1	-1							.3								-				14			
the late or or others.		-		,				,	-															
the late error whose	- 1	-	- 1	- 1			- 4		-	-		- 1					- 4	- 1	- 4		- 1			
Constitution of the consti		-1																,						

Correlation matrix, 3 of 3

Correlation Matrix

		To what extent do you feel that physical pain prevents you from doing what you need to do?	How much do you need any medical treatment to function in your daily life?	How much do you enjoy life?	To what extent do you feel your life to be meaningful?
Correlation	To what extent do you feel that physical pain prevents you from doing what you need to do?	1.0	.3	.4	.4
	How much do you need any medical treatment to function in your daily life?	.3	1.0	.2	.0
	How much do you enjoy life?	.4	.2	1.0	.8
	To what extent do you feel your life to be meaningful?	.4	.0	.8	1.0

Communalities

Communalities

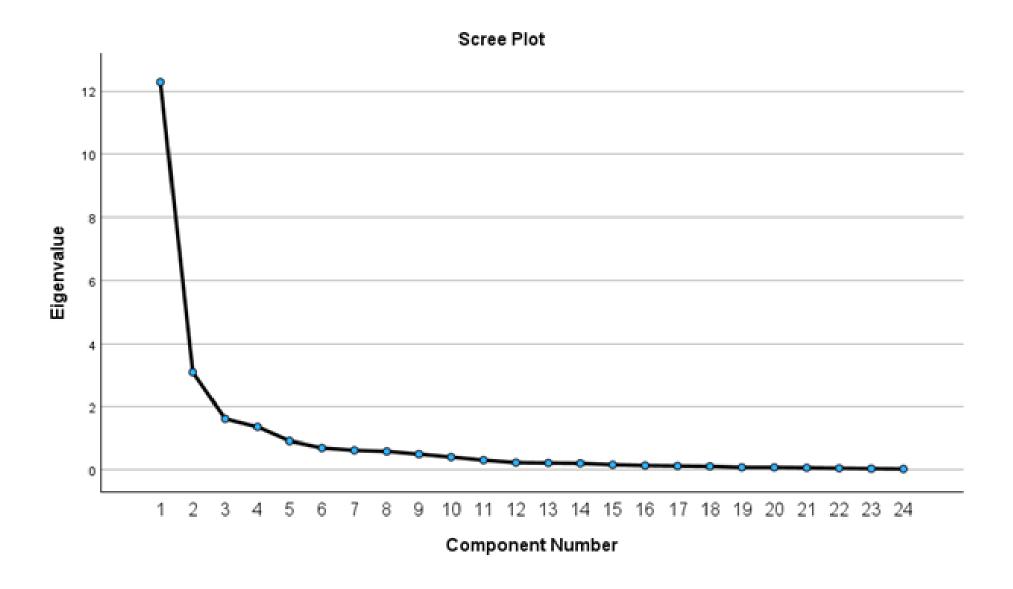
	Initial	Extraction
To what extent do you feel that physical pain prevents you from doing what you need to do?	1.000	.496
How much do you need any medical treatment to function in your daily life?	1.000	.607
How much do you enjoy life?	1.000	.814
To what extent do you feel your life to be meaningful?	1.000	.774
How well are you able to concentrate?	1.000	.797
How safe do you feel in your daily life?	1.000	.764
How healthy is your	1 000	699

Eigenvalues

Total Variance Explained

		Initial Eigenvalu	Extraction Sums of Squared Loadings						
Component	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %			
1	12.285	51.185	51.185	12.285	51.185	51.185			
2	3.101	12.923	64.108	3.101	12.923	64.108			
3	1.623	6.762	70.870	1.623	6.762	70.870			
4	1.366	5.691	76.561	1.366	5.691	76.561			
5	.918	3.824	80.385						
6	.701	2.919	83.304						
7	.627	2.612	85.915						
8	.592	2.465	88.380						
9	.505	2.103	90.483						
10	410	1 710	02 102						

Scree plot



Component matrix

Component Matrix^a

Component

	1	2	3	4
To what extent do you feel that physical pain prevents you from doing what you need to do?	.282	.557	.199	.257
How much do you need any medical treatment to function in your daily life?	073	.642	130	.415
How much do you enjoy life?	.770	.257	.376	.116
To what extent do you feel your life to be meaningful?	.745	.206	.402	.120
How well are you able to concentrate?	.852	.155	.174	.132
How safe do you feel in	.743	.125	.395	.203

Live demo, Principal components analysis

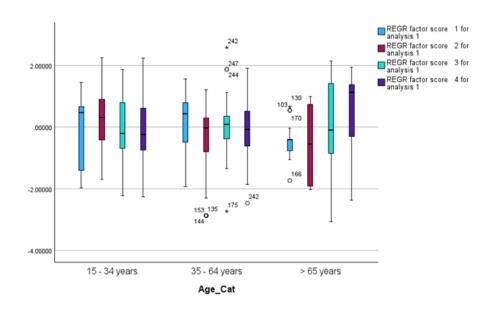
Break #1

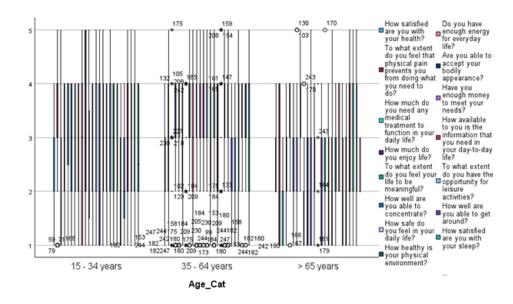
- What you have learned
 - Principal components analysis
- What's coming next
 - Applications of principal components

Applications

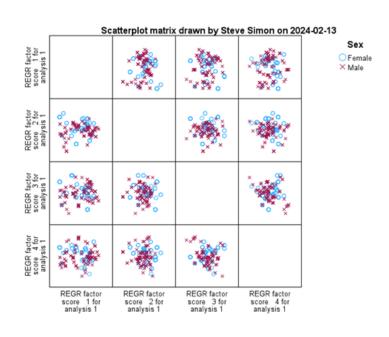
- Visualization
 - Reduce high dimensional visualization
 - Fewer graphs
- Regression analysis
 - Fewer independent variables (rule of 15)
 - Removes collinearity

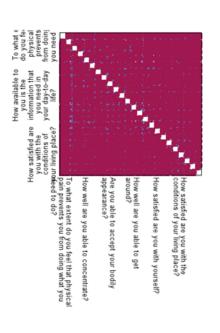
Boxplots of first four principal components





Scatterplot of first four principal components





Sex

O Female

R-squared using four principal components

Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.785ª	.617	.610	.528

a. Predictors: (Constant), REGR factor score 4 for analysis 1, REGR factor score 3 for analysis 1, REGR factor score 2 for analysis 1, REGR factor score 1 for analysis 1

R-squared using all 24 variables

Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.841 ^a	.707	.676	.482

a. Predictors: (Constant), How often do you have negative feelings such as blue mood, despair, anxiety, depression?, How much do you need any medical treatment to function in your daily life?, To what extent do you feel that physical pain prevents you from doing what you need to do?, How healthy is your physical environment?, Do you have enough energy for everyday life?, To what extent do you have the opportunity for leisure

activities?, How satisfied are you with the conditions of your living place?, How satisfied are you with your personal relationships?, How satisfied are you with the support you get from your friends?, How safe do you feel in your daily life?, Have you enough money to meet your needs?, How satisfied are you with your capacity for work?, To what extent do you feel your life to be meaningful?, How satisfied are you with your sex life?, Are you able to accept your bodily appearance?, How satisfied are you with your transport?, How much do you enjoy life?, How satisfied are you with your ability to perform your daily living activities?, How available to you is the

information that you need in

your day-to-day life?, How satisfied are you with your sleep?, How well are you able to concentrate?, How satisfied are you with your access to health services?, How well are you able to get around?

Live demo, Applications of principal components

Break #2

- What you have learned
 - Applications of principal components
- What's coming next
 - Factor analysis

Philosophy behind factor analysis

- Variance equals information
- Covariance (correlation) equals shared information
- Modeling shared information creates latent variables

Speaker notes

Most of the time, statisticians talk about variation as if it were a bad thing. It's noise. It is what makes our confidence intervals so wide. It's what forces us to collect such large sample sizes.

But, in a different context, variation is a good thing. It provides valuable information.

Consider studies of student success done at Harvard and at Johnson County Community College. There is relatively little variation at Harvard in things like high school GPA and SAT scores because everyone is clustered near the upper limits of these scales. At Johnson County Community College—a very good school, don't get me wrong, there is a bit more variation. So you can see enough of a spread to better understand factors for success after graduation.

Some clinical trials deliberately focus on the most seriously ill patients. Mildly ill patients don't have a lot of variation because there is only a small amount of improvement needed to get them back to full health. Serioulsly ill patients have more variation, giving you more opportunity to identify successful treatment options. Now, there are exceptions. If the patients are so seriously ill that no one recovers, you fall back into the same problem of too little variation.

If variance equals information, then covariance (or equivalently correlation) equals shared information.

Factor rotation

- Recombine factors
- Strive for simple interpretation
 - Components close to -1, 0, or 1
 - Each variable has one and only one non-zero components
 - Not always achievable

Rotated factor pattern, 1 of 3

Rotated Factor Matrix^a

H-1	100	grave.	T/O	E.
	100	No.	w	ш

	1	2	3	4
To what extent do you feel that physical pain prevents you from doing what you need to do?	.278	.351	066	185
How much do you need any medical treatment to function in your daily life?	.083	.160	133	510
How much do you enjoy life?	.464	.731	.108	.156
To what extent do you feel your life to be meaningful?	.391	.789	.131	.122
How well are you able to concentrate?	.522	.696	.307	.074
How safe do you feel in	.387	.693	.266	.131

Rotated factor pattern, 2 of 3

Rotated Factor Matrix^a

.810

.778

How satisfied are you with

your ability to perform your

How satisfied are you with

daily living activities?

your capacity for work?

get

around?

your sleep?

yourself?

How well are you able to

How satisfied are you with

How satisfied are you with

Are you able to accept your

1	2	3	4
.911	.276	.096	.047
.843	.274	.139	.077
.819	.396	.170	.159
.813	.293	.196	.058

.230

.090

.153

.287

.338

.340

Factor

Rotated factor pattern, 3 of 3

Rotated Factor Matrix^a

	Factor			
	1	2	3	4
How satisfied are you with your ability to perform your daily living activities?	.911			
How satisfied are you with your capacity for work?	.843			
How well are you able to get around?	.819			
How satisfied are you with your sleep?	.813			
How satisfied are you with yourself?	.810			
Are you able to accept your	.778			

Live demo, Factor analysis

Break #3

- What you have learned
 - Factor analysis
- What's coming next
 - Criticisms of principal components analysis and factor analysis

Criticisms of principal components analysis

- Advantages
 - Makes collection of many variables manageable
 - Eliminates collinearity issues
 - Focus only on important sources of variation
- Disadvantages
 - Components often uninterpretable
 - False sense of parsimony



Criticisms of factor analysis

- Advantages
 - Explore underlying structure
 - Create or validate subscales
- Disadvantages
 - Difficulty in choosing number of factors
 - Reification



Summary

- What you have learned
 - Principal components analysis
 - Applications of principal components
 - Factor analysis
 - Criticisms of principal components analysis and factor analysis

Additional topics??

Speaker notes

- Learning objectives
 - To be determined