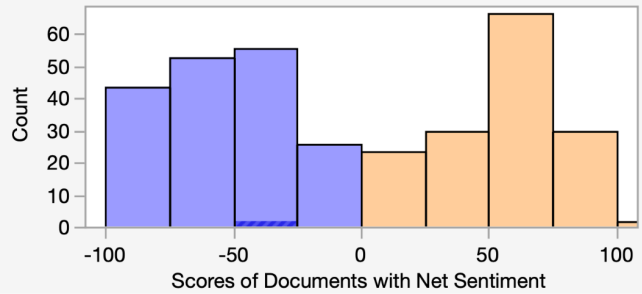
Sentiment Analysis

2.2.1.1 Trustpilot dataset

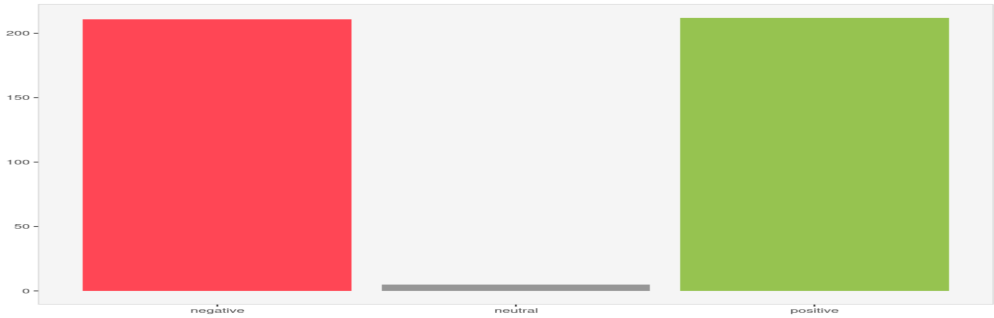
Sentiment Summary

	N	Mean Score
All Scored Documents	326	-3.0
Net Positive Documents	148	54.5
Net Negative Documents	175	-54.2
No Sentiment Documents	102	0.0

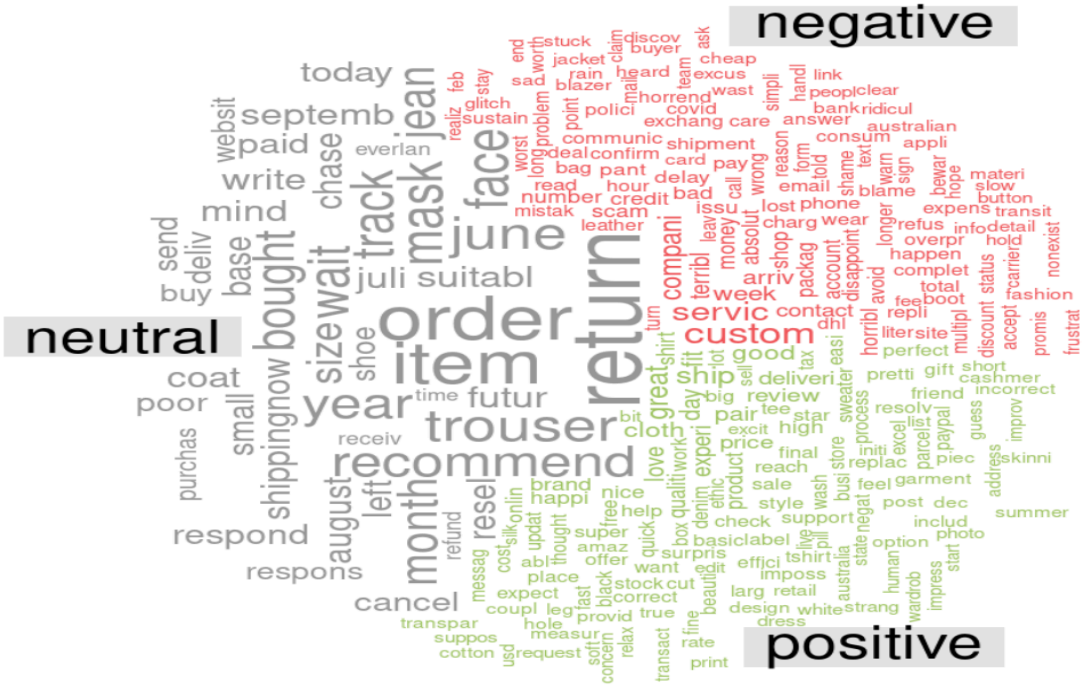


JMP sentiment summary

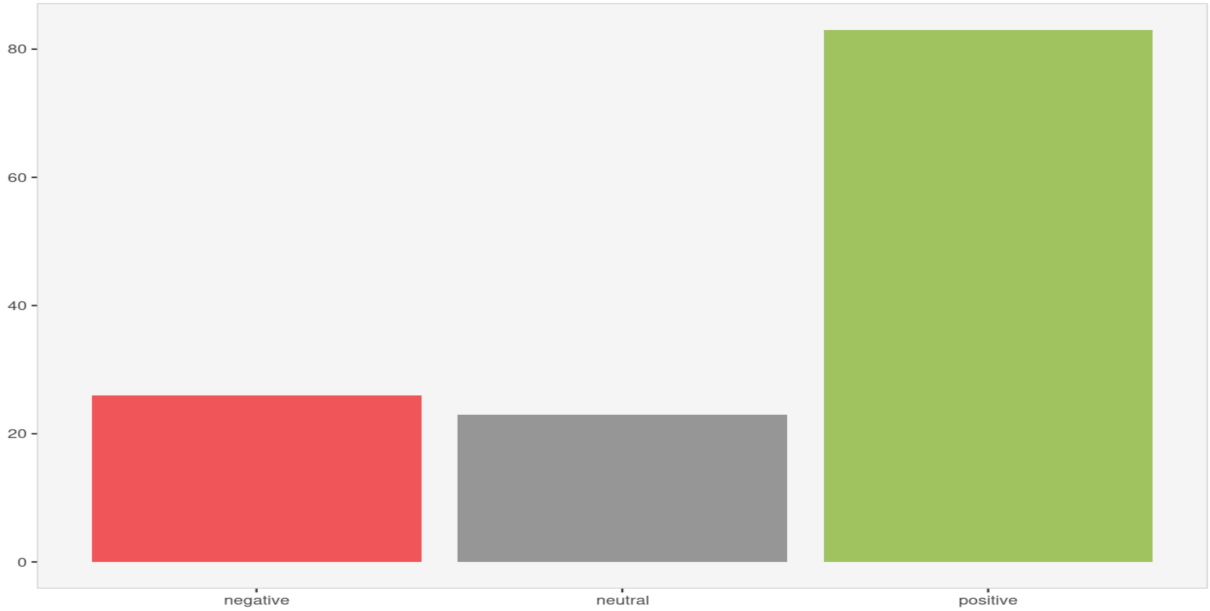
Word cloud after stemming



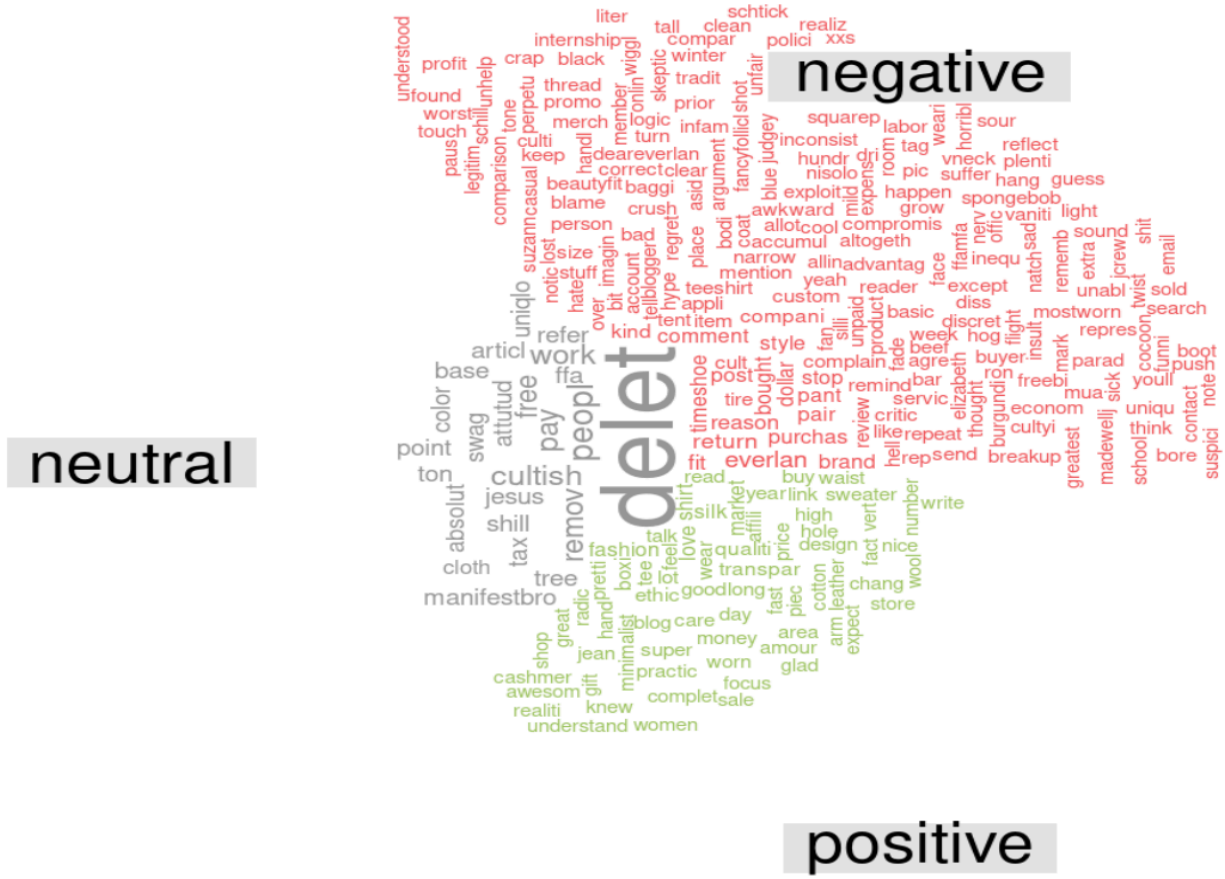
Enginius sentiment classification


negative
neutral
positive
return
order
item
recommend
trouser
year
june
mask
face
track
wait
bought
size
month
jean
today
septemb
paid
write
mind
send
deliv
buy
base
chase
everlan
juli
suitabl
futur
time
coat
poor
small
shippingnow
purchas
respond
respons
cancel
august
left
resel
refund
shoe
receiv
custom
servic
contact
ship
great
cloth
price
review
good
deliveri
week
arriv
company
terribl
money
absolut
charg
wear
lost
phone
issu
scam
mistak
credit
bad
delay
pay
card
confirm
deal
communic
shipment
exchang
care
answer
consum
appli
australian
covid
polici
sustain
glitch
problem
point
worst
long
reason
form
told
shame
text
warm
sign
bewar
hope
blame
email
hour
read
number
bag
pant
leather
happen
complet
total
overpr
refus
info
detail
hold
button
slow
materi
expens
transit
longer
avoid
account
disappoint
packag
shop
fashion
boot
fee
repli
horribl
liter
site
perfect
pretti
gift
short
cashmer
incorrect
friend
resolv
guess
improv
paypal
excel
list
parcel
piec
skinni
address
summer
dec
post
includ
photo
option
start
impress
wardrob
human
australia
state
negat
support
check
style
sale
feel
garment
reach
final
replac
store
busi
wash
process
sweater
easi
tax
tot
sell
big
day
work
experi
product
ethic
denim
qualiti
love
box
quick
free
help
nice
happi
brand
updat
onlin
silk
cost
messag
thought
amaz
offer
want
edit
effici
tshirt
live
pill
surpris
basic
label
place
abl
black
fast
stock
cut
correct
beautif
imposs
larg
retail
design
white
strang
dress
true
provid
leg
coupl
expect
hole
transpar
suppos
measur
cotton
usd
request
soft
concern
relax
transact
fine
rate
print
turn
shirt
tot
fit
day
bit
pair
tee
star
excit
high
final
reach
sale
style
check
support
basic
label
live
pill
state
australia
human
wardrob
impress
start
option
photo
includ
post
dec
summer
address
skinni
piec
paypal
guess
improv
incorrect
cashmer
short
gift
pretti
perfect
site
liter
multipl
discount
status
accept
promis
frustrat
nonexist
carrier
fashion
boot
fee
repli
horribl
avoid
complet
total
happen
overpr
refus
info
detail
hold
button
slow
materi
expens
transit
longer
account
disappoint
packag
shop
wear
charg

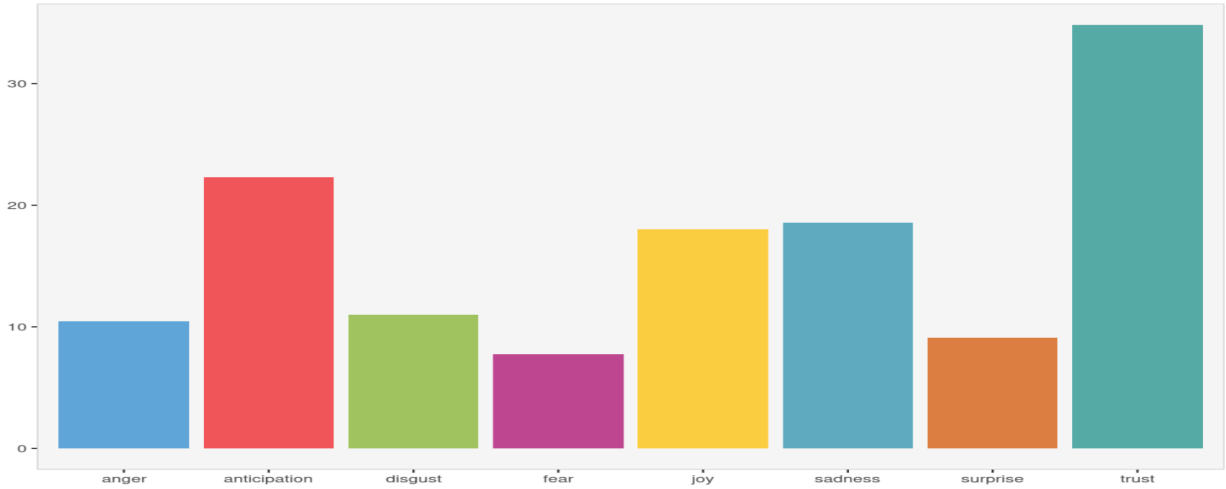
disgust
anticipation
fear
anger
joy
trust
sadness
surprise
great
avoid
qualiti
love
product
shoe
boot
costterribl
absolut
item
order
time
receivemail
servic
disappoint
everlan
charg
custom
jean
easi
good
basic
tax
size
credit
return
shirt
cloth
repli
dec
wash
sandal
offer
medium
amaz
nonexist
impress
negat
contact
multipl
support
surpris
shipping
week
long
deliveri
cancel
discount
pretti
trash
read
develop
sustain
accept
want
complet
construct
adjust
mediocr
expens
slow
horribl
horrend
slipper
ridicul
huge
carrier
buyer
stick
fact
mail
easili
shapeless
employe
modern
union
rubbish
dread
switzerland
doorstep
stitch
enter
bought
poshmark
difficult
exchang
fraudul
deserv
boss
theft
stain
polish
answer
add
reach
mule
reason
fur
step
everlast
bot
blah
glitch
fraud
poor
short
address
look
pay
stock
blame
tshirt
purchas
excus
told
refund
issu
recommend
number
communic
store
bank
respons
card
brand
label
account
package
final
ship
cut
mouth
debit
pant
arriv
wait
parcel
track
check
thought
buy
tee
today
follow
singl
unauthor
dishonest
restock
deserv
fraudul
exchang
link
market
wore
joy
appal
question
classic
heel
base
retail
favorit
year
money
idea
ethic
help
transpar
wear
resolut
post
sell
worth
updat
unbeliev
jacket
entir
face
thiev
earli
chino
happi
pair
coupl
dress
colour
black
eye
star
excel
cashmer
process
price
sale
fan
busi
free
piec
leav
turn
bag
industri
item
feel
care
aso
develop
web
sit
week
read
mouth
cut
final
ship
time
email
servic
purchas
tshirt
excus
pay
address
everlan
custom
label
brand
card
refund
recommend
number
communic
store
bank
respons
credit
return
shirt
size
order
account
package
tax
avoid
great
qualiti
love
product
shoe
boot
costterribl
absolut
item
good
basic
sale
fan
process
price
leav
turn
excel
cashmer
big
red
fit
bad
review
sweater
total
worn
lost
deliv
beauti
wardrob
high
denim
dutiproblem
shame
status
garment
dirt
super
friend
clear
outfit
comfort
model
courier
photo
actual
polici
top
dec
wash
sandal
tankfast
offer
medium
extra
drop
amaz
live
shop
nonexist
impress
negat
contact
multipl
support
surpris
sad
quick
slip
tote
send
decent
june
strang
larg
worst
log
notific
awesom
marin
experi
nice
guess
scam
xmas
repli
cloth
brief
problem
shame
status
courier
photo
actual
polici
dirt
garment
deliv
worn
lost
nice
guess
scam
xmas
repli
cloth
brief
worst
log
notific
awesom
marin
experi
strang
larg
decent
june
send
tote
slip
quick
surpris
support
sad
contact
multipl
nonexist
impress
negat
amaz
live
shop
offer
medium
extra
drop
tankfast
sandal
dec
wash
shame
status
courier
photo
actual
polici
top
clear
outfit
comfort
model
super
friend
denim
dutiproblem
shame
status
garment
dirt



Enginius sentiment classification



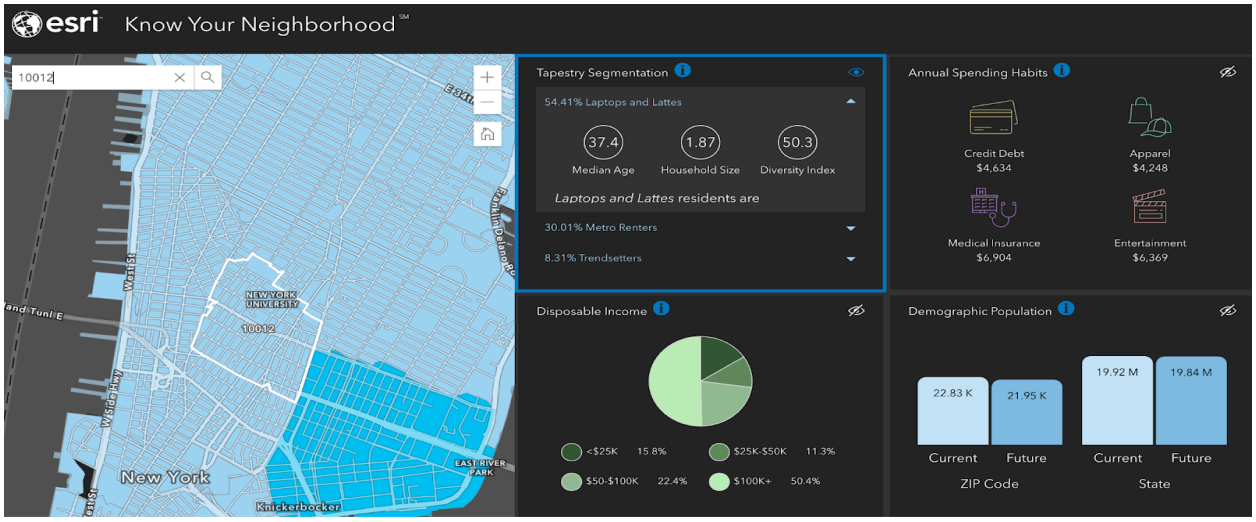
Word cloud (negative and positive)

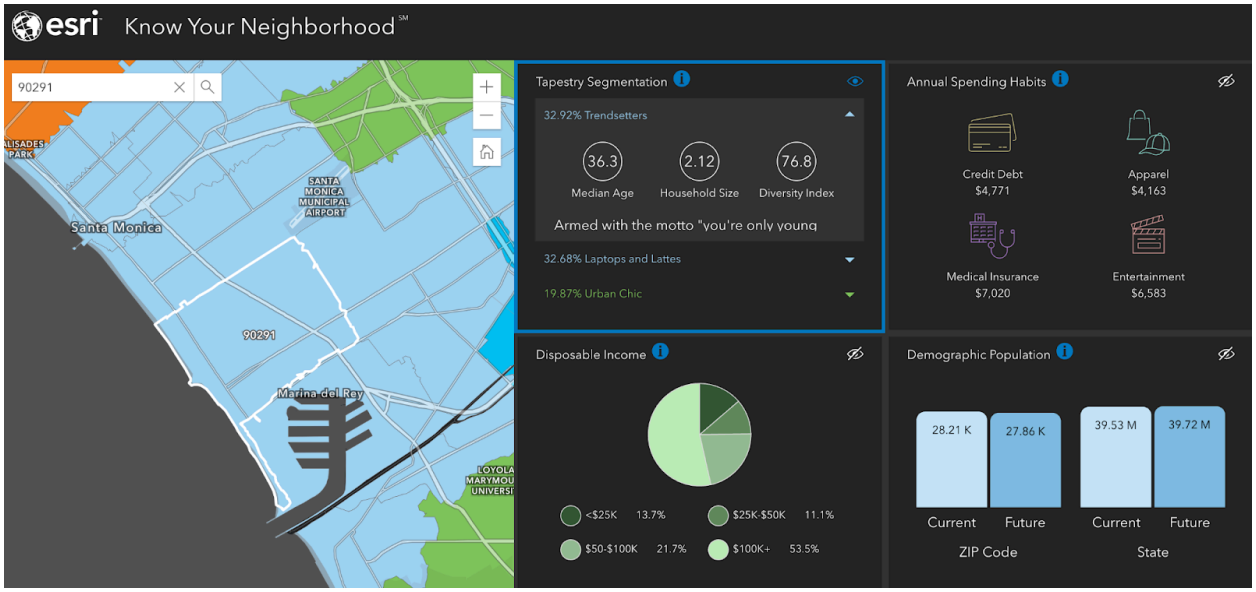


Emotion analysis (bar plots)

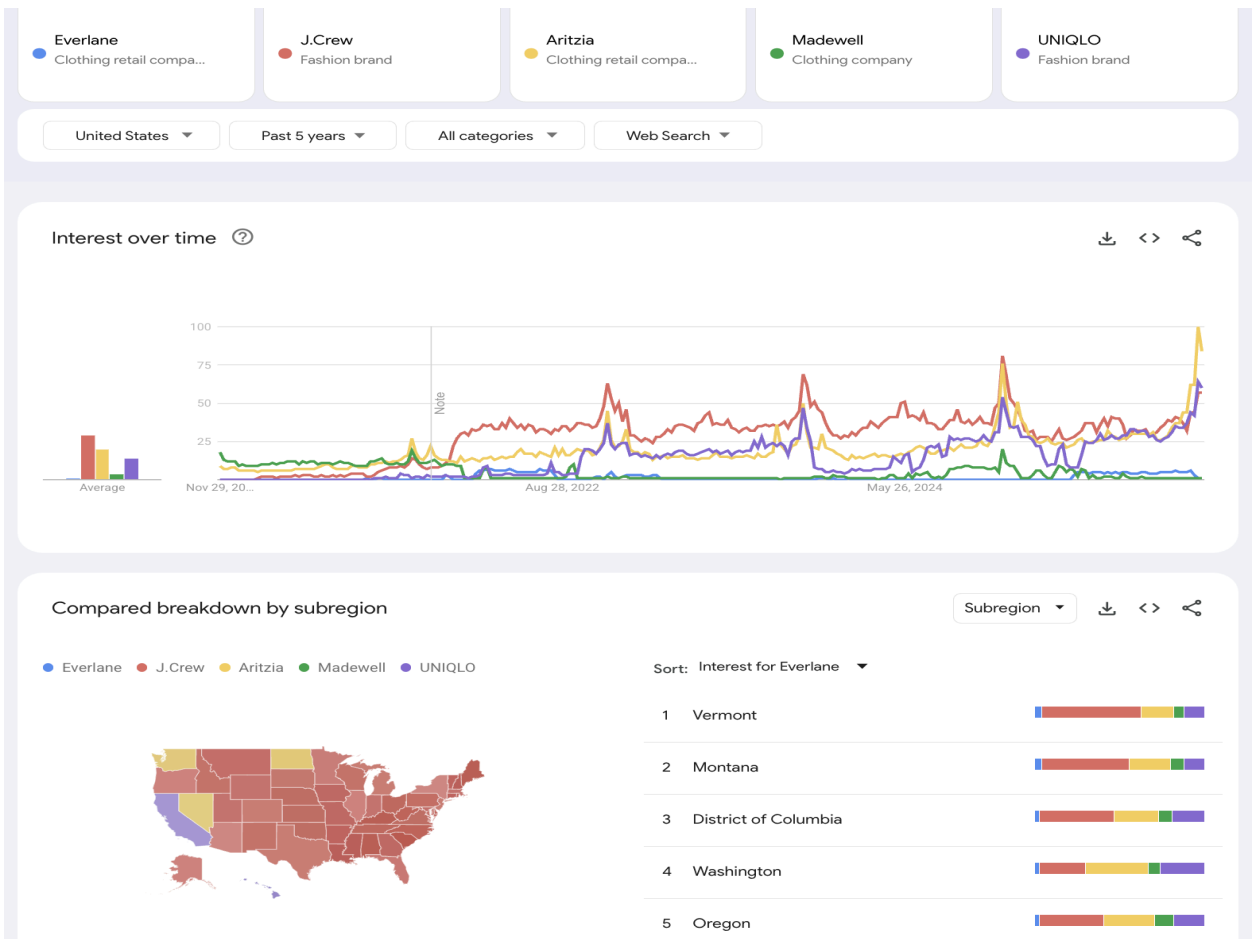
Word cloud (Emotion)

2.3 ESRI Tapestry





2.5 Google Trends



Everlane
Clothing retail compa...

ayr
Search term

Alex Mill
Topic

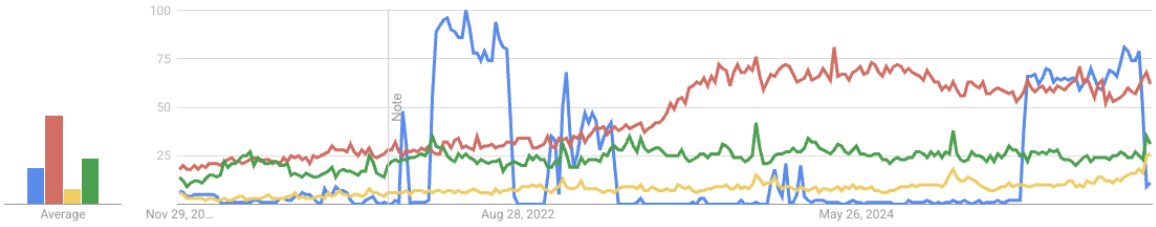
agolde
Search term

+

United States | Past 5 years | All categories | Web Search

Note: This comparison contains both Search terms and Topics, which are measured differently. LEARN MORE

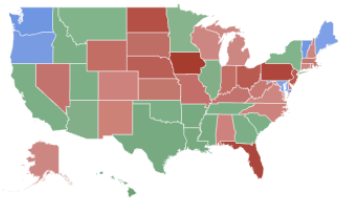
Interest over time



Compared breakdown by subregion

Subregion

Everlane ayr Alex Mill agolde



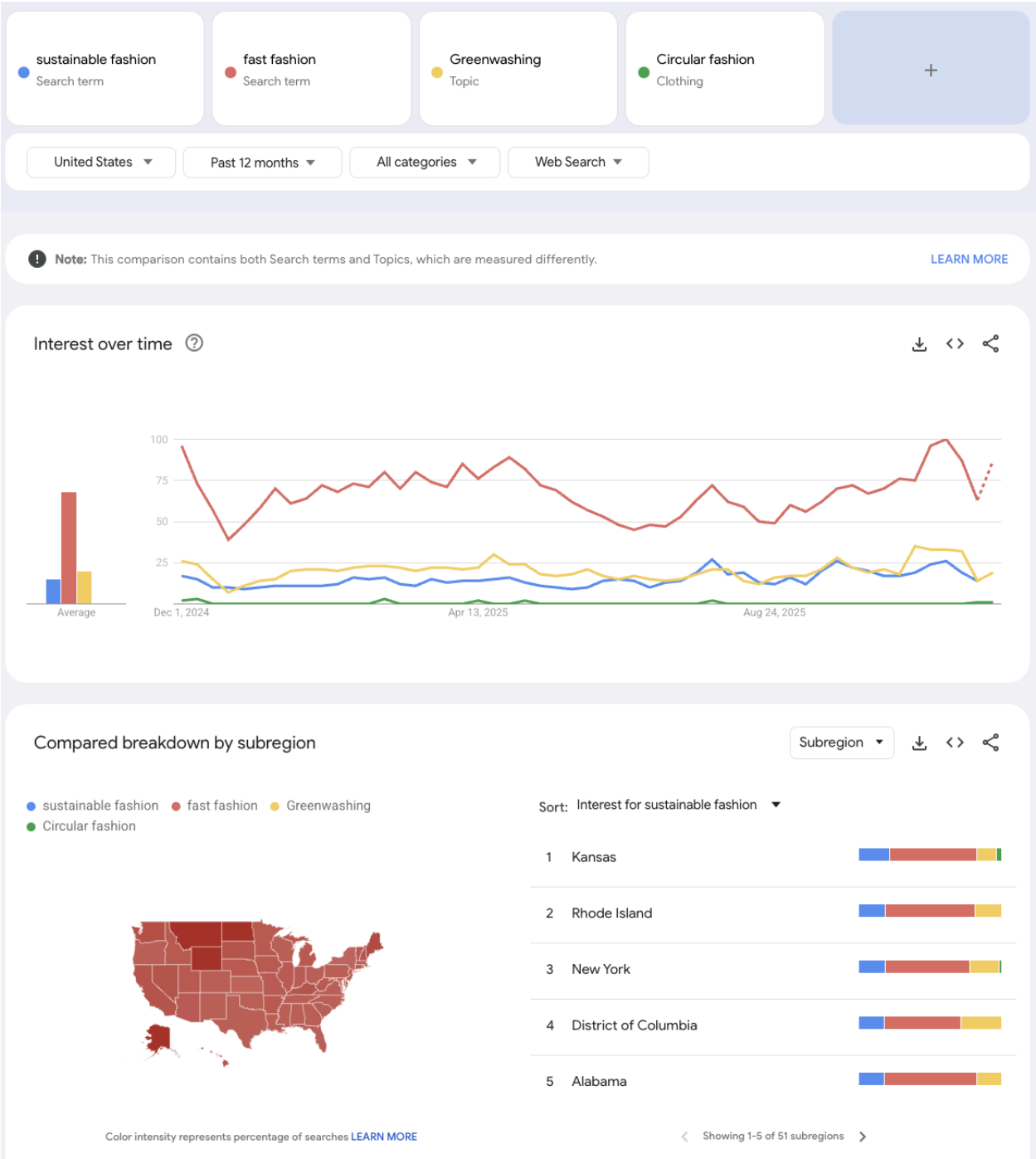
Color intensity represents percentage of searches LEARN MORE

Sort: Interest for Everlane

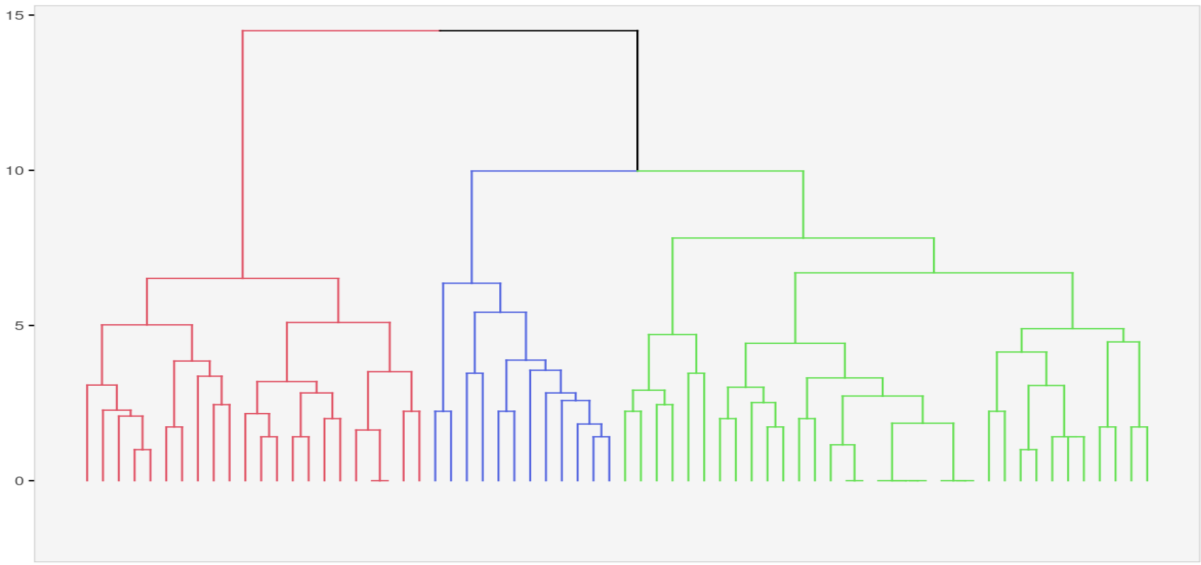
1	Washington
2	District of Columbia
3	Vermont
4	Oregon
5	California

Showing 1-5 of 51 subregions

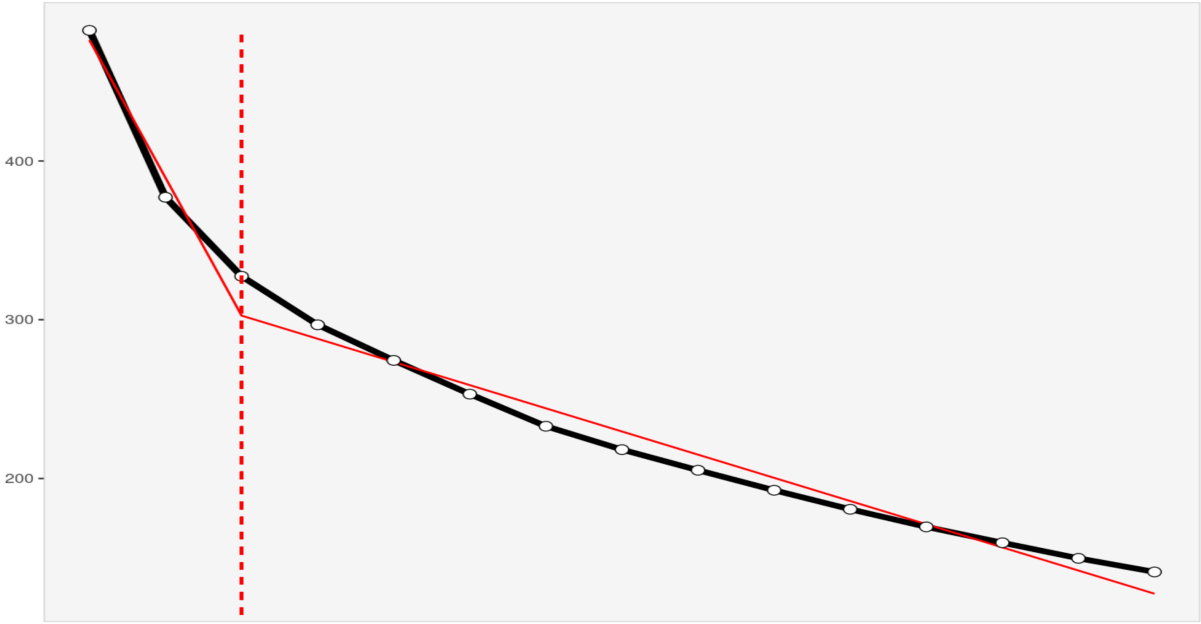
Relevant search terms centered around Everlane:



3.2 Segmentation Analytics

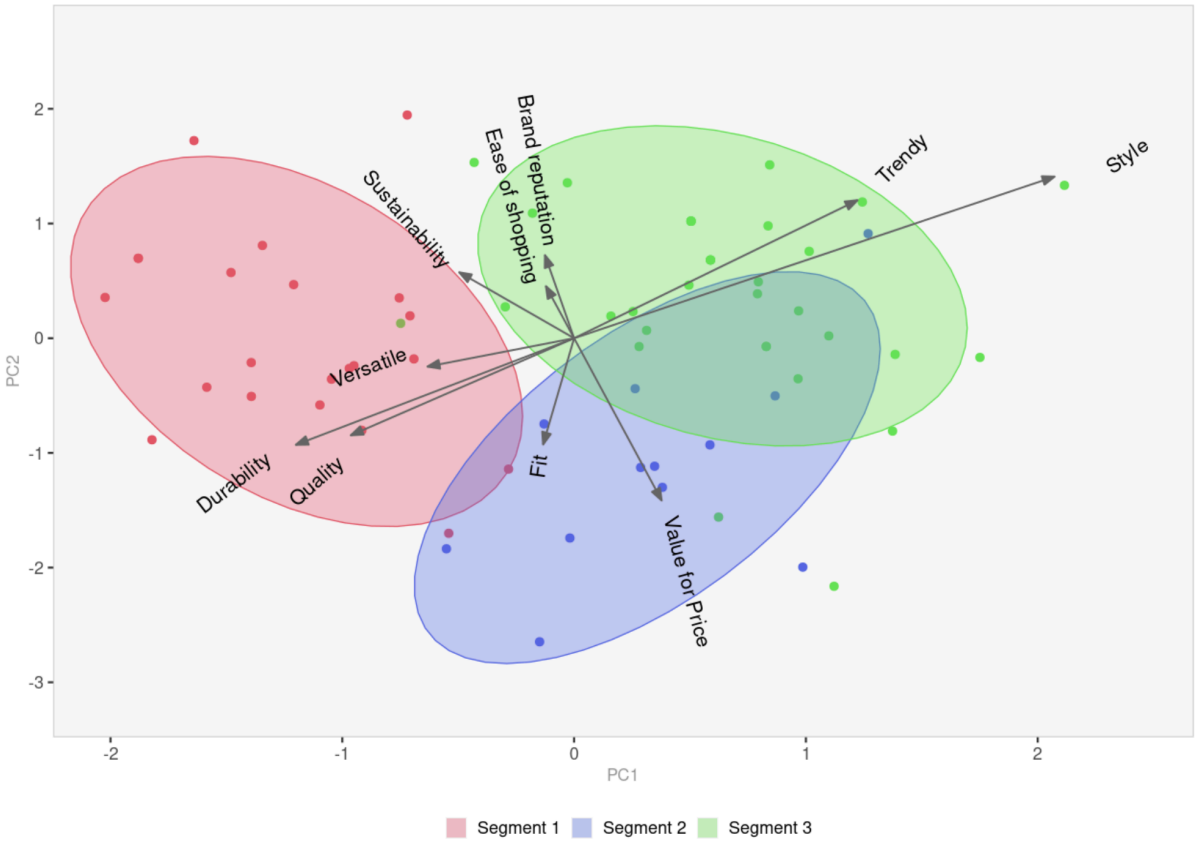


Dendrogram



Scree Plot

Segment differences per segment



Segment clusters

Descriptor analysis

Descriptors

This table reports the descriptor averages of each segment. The more differences can be found, the easier it will be to predict segment membership based on descriptors alone.

	Population	Segment 1	Segment 2	Segment 3
`1. Are you familiar with the brand Everlane?`	0.353	0.227	0.333	0.441
`2. Have you ever purchased from Everlane?`	0.191	0.091	0.25	0.235
`3. How aware are you of Everlane's sustainability or transparency initiatives (e.g., recycled materials, ethical factories)?`	1.6	1.5	2.17	1.47
`4. In the past, how much have Everlane's sustainability claims influenced your likelihood of purchasing from them?`	1.79	1.64	1.83	1.88
`5. How much will Everlane's sustainability claims influence your likelihood of purchasing from them in the future?`	2.25	2.09	2.25	2.35
`6. How credible do you find Everlane's sustainability and ethical manufacturing messaging?`	2.46	2.32	2.25	2.62
`7. How important are sustainability and ethical practices to you when choosing a clothing brand?`	2.94	3.09	2.75	2.91
`8. How well does Everlane's actual performance match its stated environmental goals?`	2.65	2.68	2.5	2.68
`9. Compared to other clothing brands, how strong do you feel Everlane's sustainability efforts are?`	2.93	2.64	3.5	2.91
`10. How easy is it to find and purchase Everlane	3.07	2.68	3.17	3.29

products (online or in stores)?`				
`11. How satisfied are you with Everlane's online shopping experience?`	3.15	2.64	3.42	3.38
`12. How easy is it to return or exchange Everlane items?`	3.24	2.95	3.58	3.29
`13. How well does Everlane's pricing align with the quality offered?`	3	2.55	3	3.29
`14. How likely are you to recommend Everlane to a friend or family member?`	2.84	2.82	3.08	2.76
`15. How likely are you purchase Everlane in the next 6 months?`	2.63	2.5	2.58	2.74
`16. How satisfied are you with the brand Everlane as a whole?`	2.71	2.68	2.92	2.65
`1. Age range:`	1.59	1.5	1.75	1.59
`2. Gender identity:`	0.515	0.5	0.167	0.647
`3. Annual household income:`	2.35	2	2.83	2.41
`4. Education level:`	2.43	2.59	2.58	2.26
`5. Employment status:`	1.97	1.55	2.33	2.12
`6. Relationship status:`	1.49	1.36	1.75	1.47
`7. Geographic location:`	1.18	1.09	1	1.29
ShopFreq	3.49	3.23	2.58	3.97
SpendMonth	2.16	2.18	1.5	2.38
ItemsPerTrip	1.75	2	1.42	1.71
SM_DiscoverFreq	3.51	3.45	2.33	3.97

Descriptor data per segment. Average value of each descriptor, overall and within each cluster. Descriptors that are statistically different from the rest of the population are highlighted in red (lower) or green (higher).

3.3. Positioning Analytics

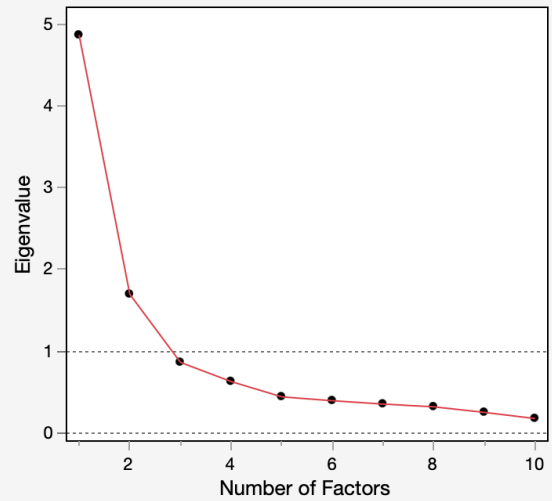
3.3.2 Factor Analysis

Factor Analysis Student Edition

Eigenvalues

Number	Eigenvalue	Percent	Cum Percent
1	4.8655	48.655	48.655
2	1.6970	16.970	65.626
3	0.8641	8.641	74.267
4	0.6297	6.297	80.564
5	0.4428	4.428	84.992
6	0.3952	3.952	88.944
7	0.3546	3.546	92.490
8	0.3209	3.209	95.698
9	0.2528	2.528	98.226
10	0.1774	1.774	100.000

Scree Plot

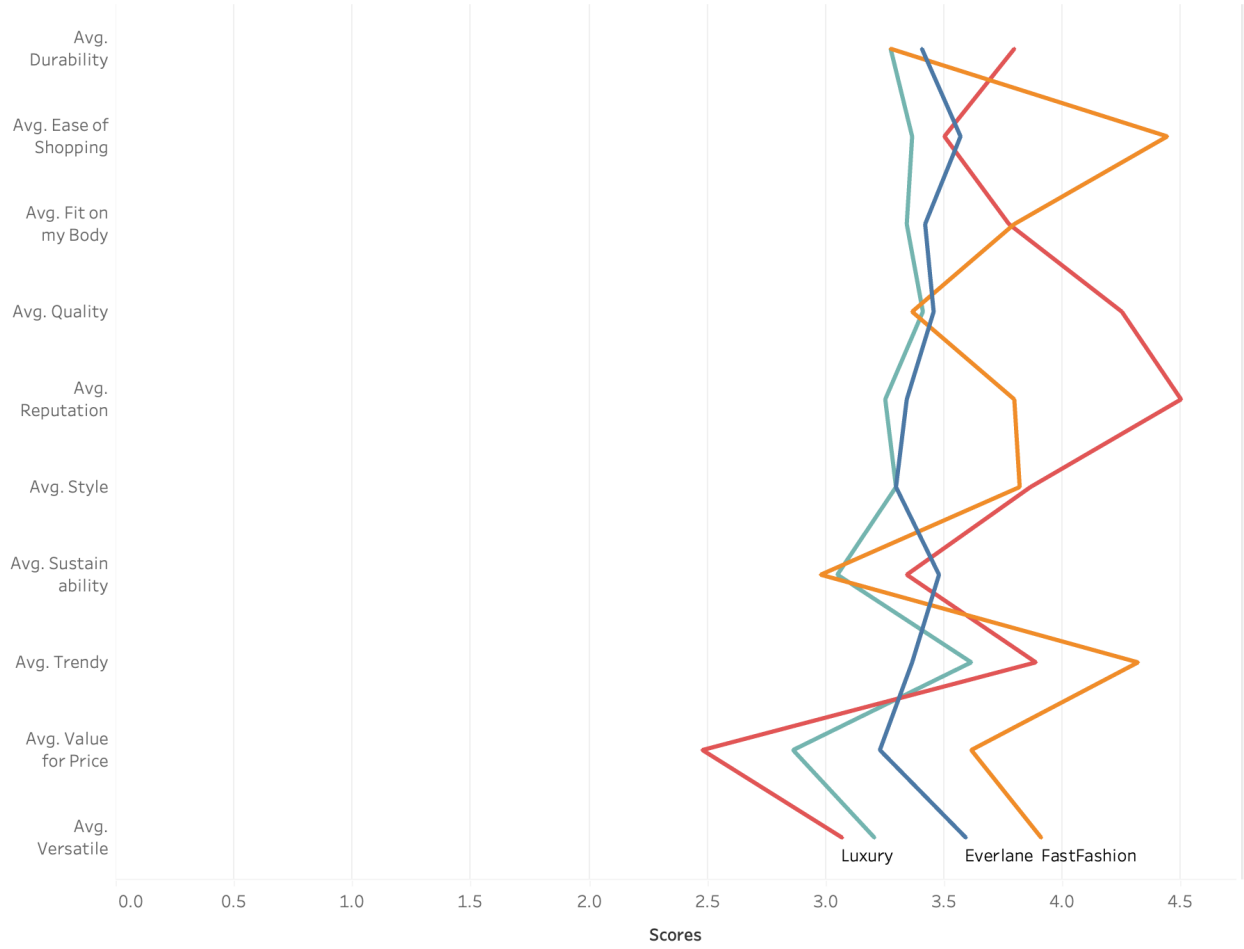


Variance Explained by Each Factor

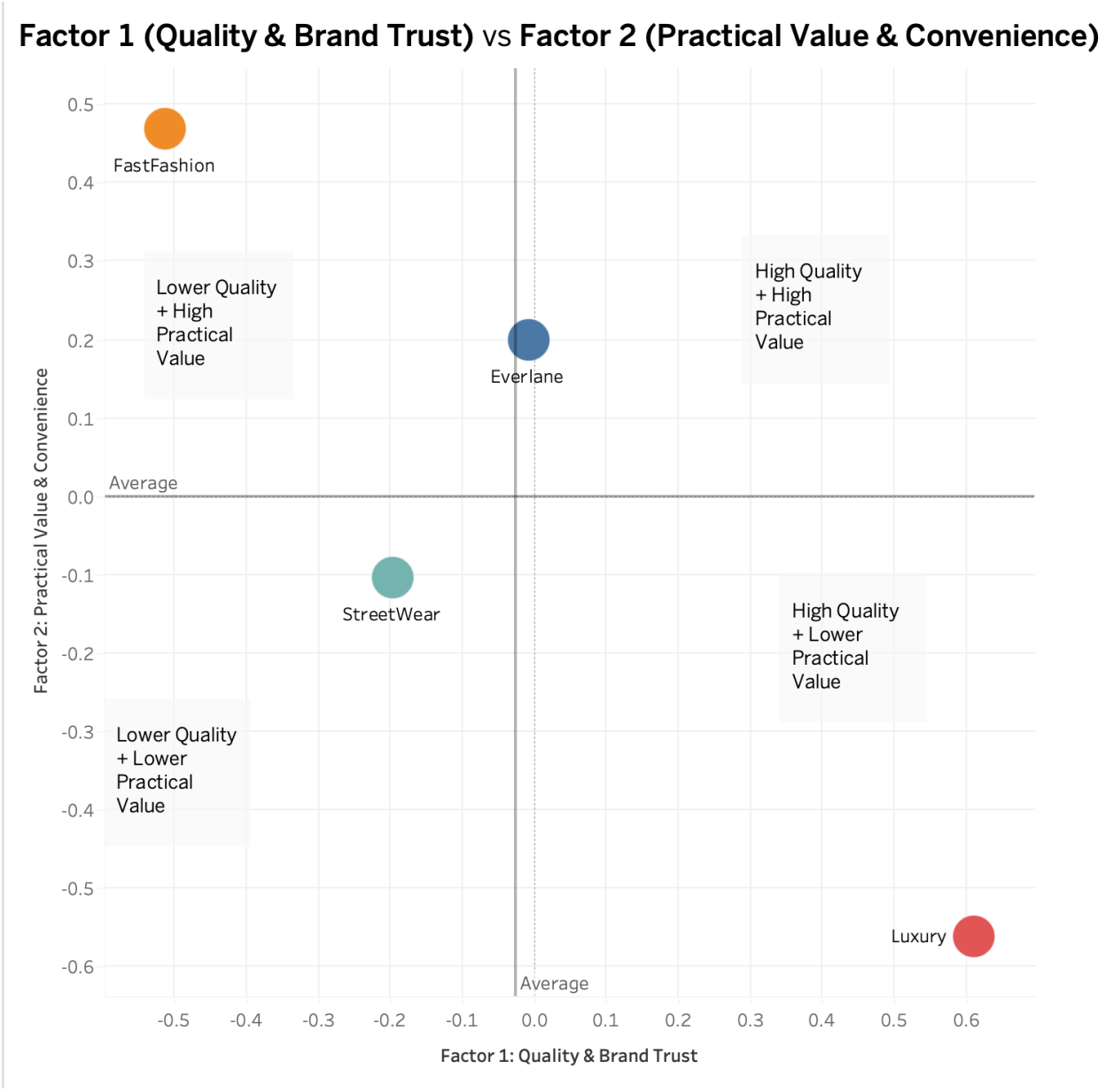
Factor	Variance	Percent	Cum Percent
Factor 1	2.8932	28.932	28.932
Factor 2	2.0153	20.153	49.085
Factor 3	1.4726	14.726	63.811

3.3.4 Attribute-Level Insights (Snake Plot)

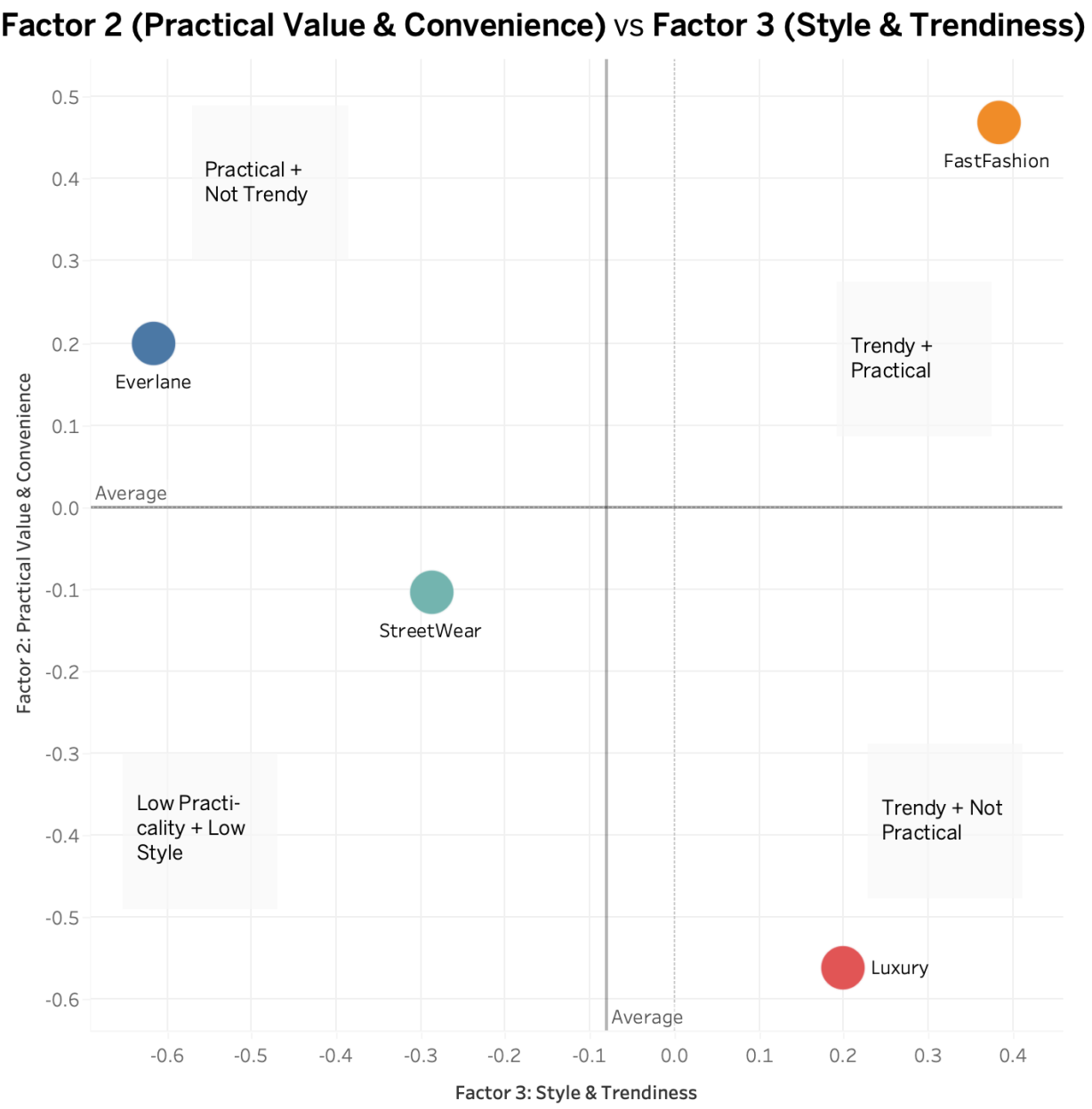
Snake Plot: Attribute-Level Brand Profiles Across 10 Dimensions



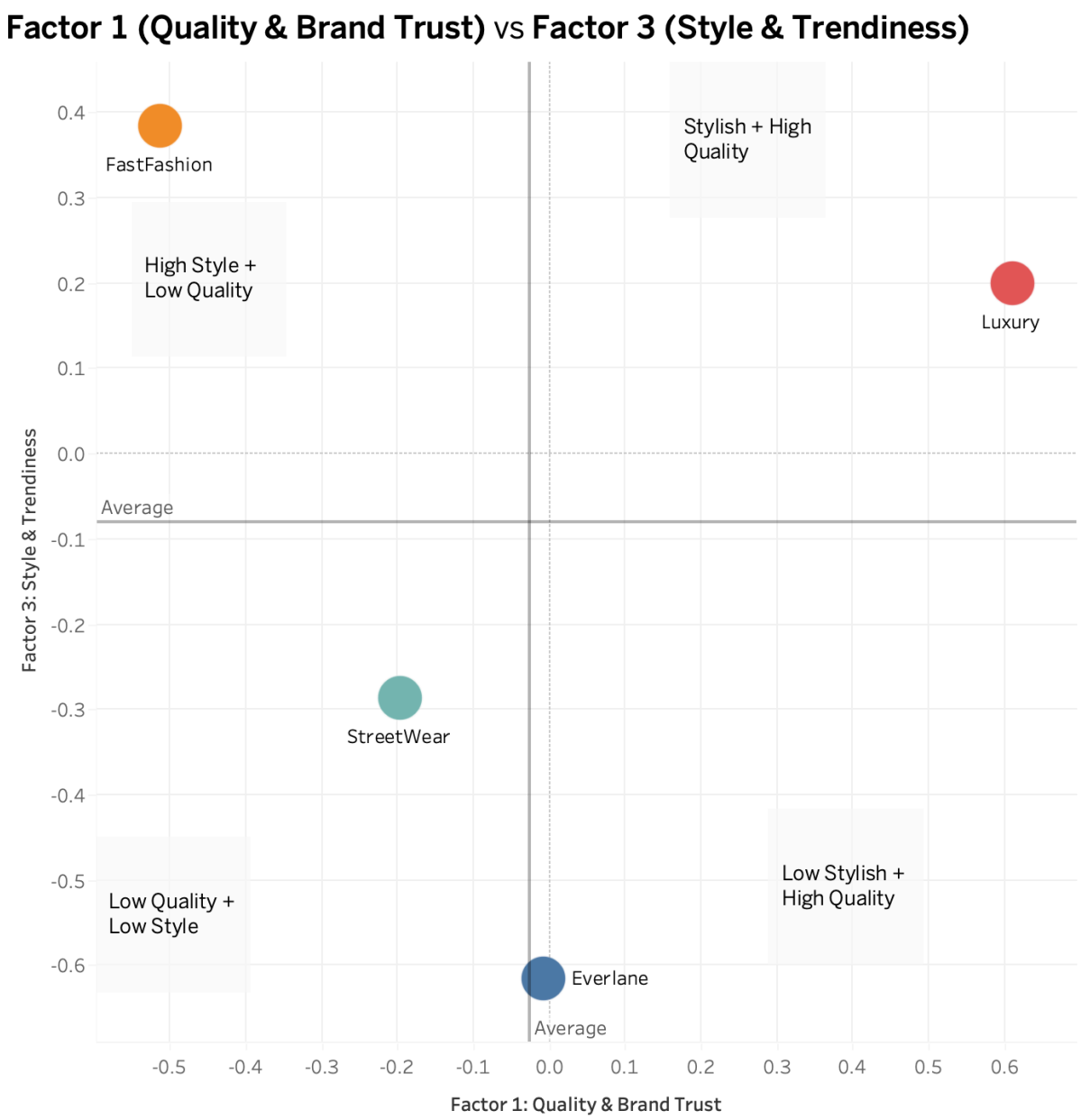
3.3.3 Perceptual Mapping Factor 1 vs Factor 2: Quality & Trust × Practical Value



3.3.3 Perceptual Mapping Factor 2 vs Factor 3: Practical Value × Style



3.3.3 Perceptual Mapping Factor 1 vs Factor 3: Quality × Style



3.4 Predictive Analytics

Nominal Logistic Fit for everlane_purchase_intent_6 months 2

Effect Summary

Source	Logworth	PValue
2. Have you ever purchased from Everlane?	4.030	0.00009
2. Gender identity:	1.559	0.02760

Converged in Gradient, 6 iterations

Whole Model Test

Model	-LogLikelihood	DF	ChiSquare	Prob>ChiSq
Difference	9.214785	2	18.42957	<.0001*
Full	20.880969			
Reduced	30.095754			

RSquare (U)	0.3062
AICc	48.1369
BIC	54.4205
Observations (or Sum Wgts)	68

Lack Of Fit

Source	DF	-LogLikelihood	ChiSquare
Lack Of Fit	1	0.257117	0.514235
Saturated	3	20.623852	Prob>ChiSq
Fitted	2	20.880969	0.4733

Parameter Estimates

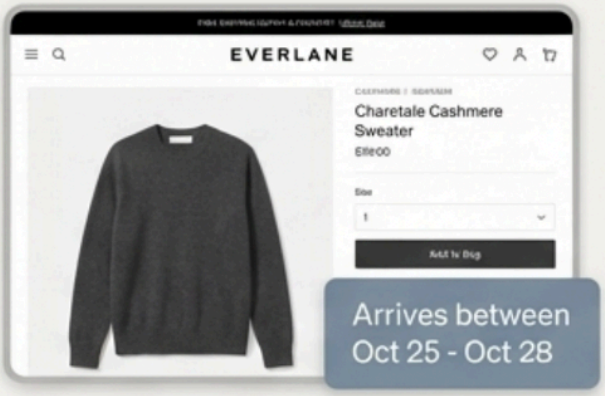
Term	Estimate	Std Error	ChiSquare	Prob>ChiSq
Intercept	-1.2999432	0.4209073	9.54	0.0020*
2. Gender identity:[Female]	-0.9296354	0.4695607	3.92	0.0477*
2. Have you ever purchased from Everlane?[No]	-1.571265	0.4530323	12.03	0.0005*

For log odds of 1/0

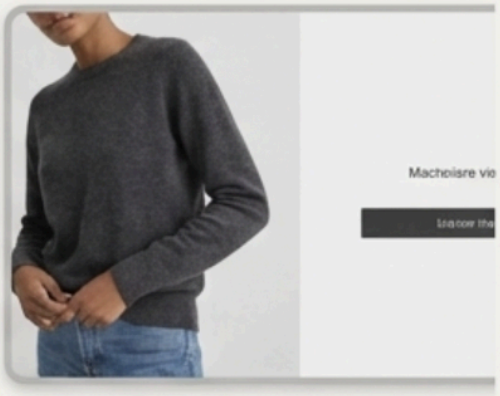
Effect Likelihood Ratio Tests

Source	Nparm	DF	L-R ChiSquare	Prob>ChiSq
2. Gender identity:	1	1	4.85297025	0.0276*
2. Have you ever purchased from Everlane?	1	1	15.2682977	<.0001*

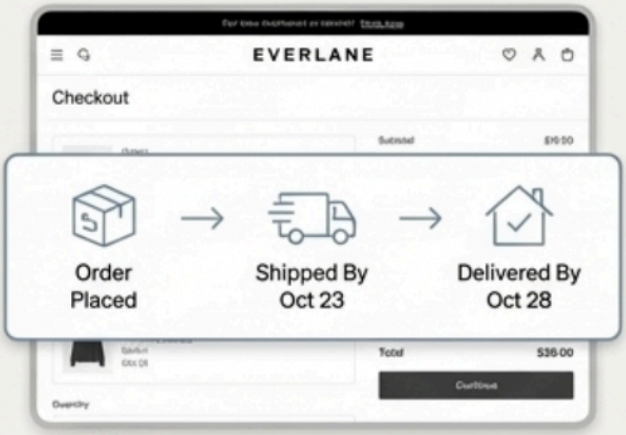
4. MOCK-Ups



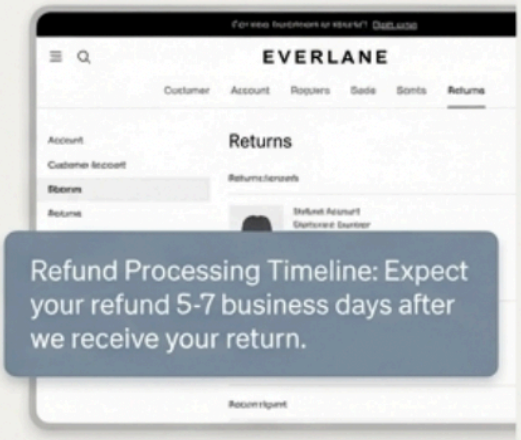
Mock-up 1 (Product Page)



Mock-up 3 (Peturnt Page)

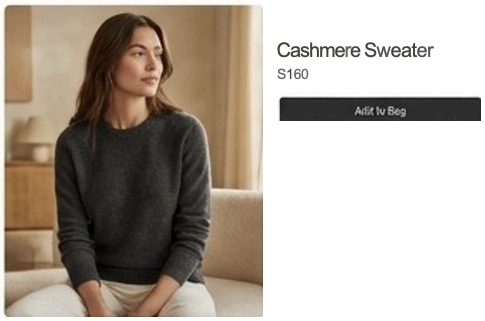


Mock-up 2 (Checkout)

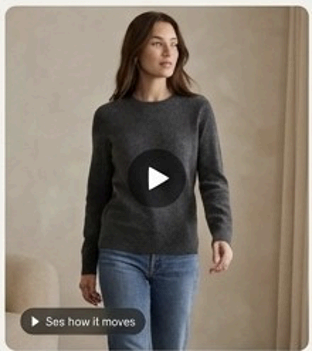


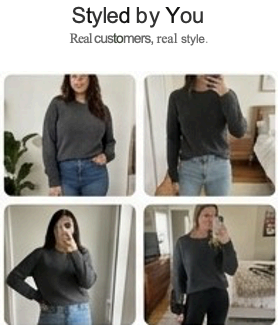
Mock-up 3 (Returns Page)

Before: Current Approach



After: Proposed 'Product Proof' Approach





Targeted Expansion
Use the ESRI "Trendsetter" profile to target lookalike ZIP codes (e.g., SoMa in SF, East Village in NYC, Wicker Park in Chicago) for digital campaigns and future retail expansion.

enginius
MARKETING ENGINEERING ONLINE

Enginius

Sentiment Analysis

Charudisha Ashjay, University of Southern California

Copyright (c) 2025, DecisionPro Inc.

Sentiment options

Options selected

	Option	Selection
	Data source	Sentiment data
	Verbatim	text
	Word co-occurrence analysis and RAKE	No
	Topic model	No
	Default stop words	Long list (851)
	Custom stop words	No
	Date and time	2025-10-21 01:45:41 UTC

Options selected.

Data description

	Data	Number of Rows	Number of columns	Column names
1	Sentiment data	428	5	doc_id, text, rating_numeric, sentiment_label_from_rating, is_english_guess

Data description.

Word cloud

Word cloud of the most frequent words

Word cloud. The word cloud represents the most frequently used words inside the corpus of texts provided. The bigger a word appears, the larger the number of times it occurs in the text corpus.

Word cloud of the most frequent words (without stemming)

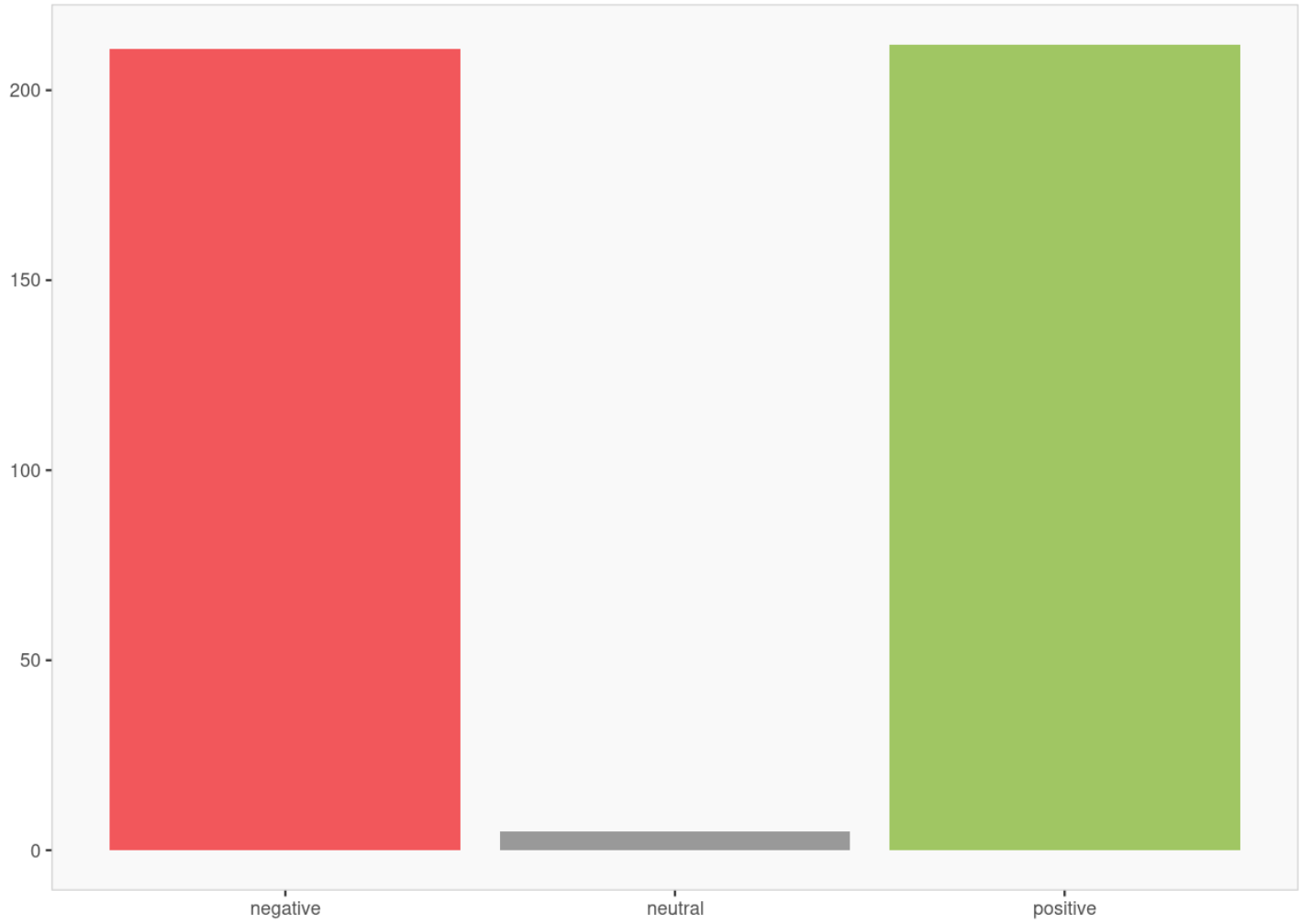
Word cloud without stemming. The word cloud represents the most frequently used words inside the corpus of texts provided. The bigger a word appears, the larger the number of times it occurs in the text corpus.

Valence analysis

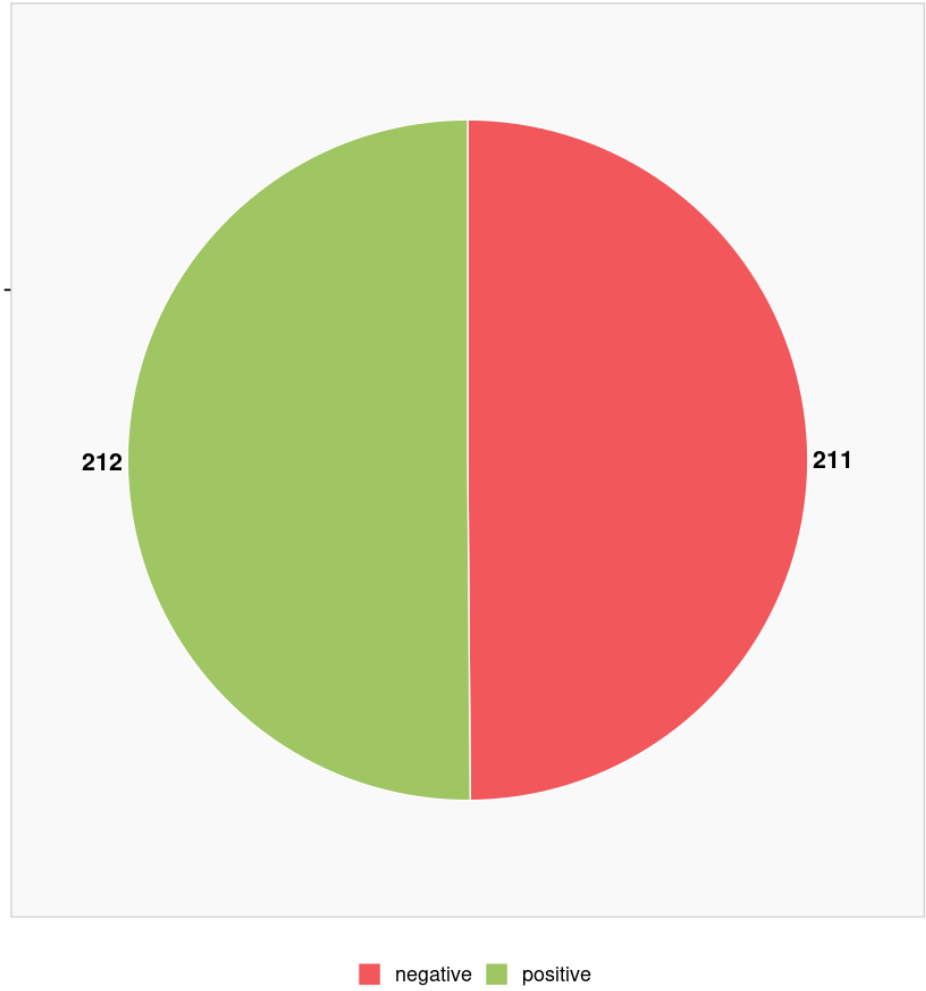
Valence repartition

	Posts count	Relative posts count
Total	428	100%
negative	211	49%
neutral	5	1%
positive	212	50%

Valence repartition. The number of posts that fall into different valence categories summarized by their absolute and relative values.

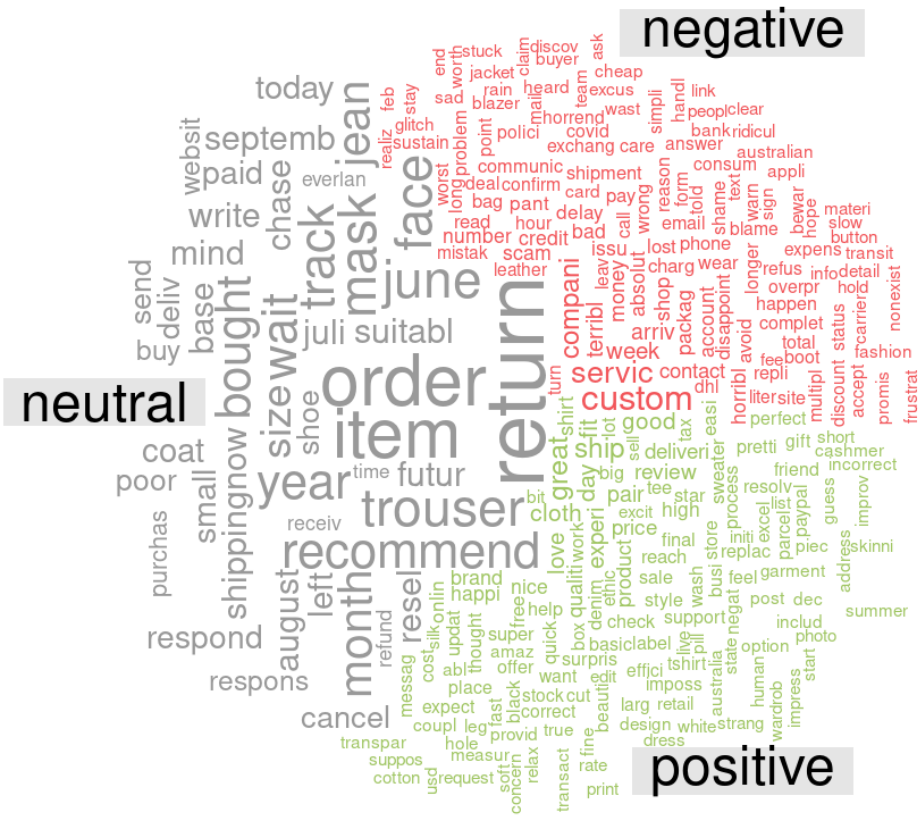


Valence histogram. The valence histogram indicates the number of posts by their valence.



Valence distribution without uncategorized posts. Graphic summary of the relative sizes of the number of posts classified by valence after ignoring posts that could not be categorized (i.e., neutral posts).

Valence word cloud

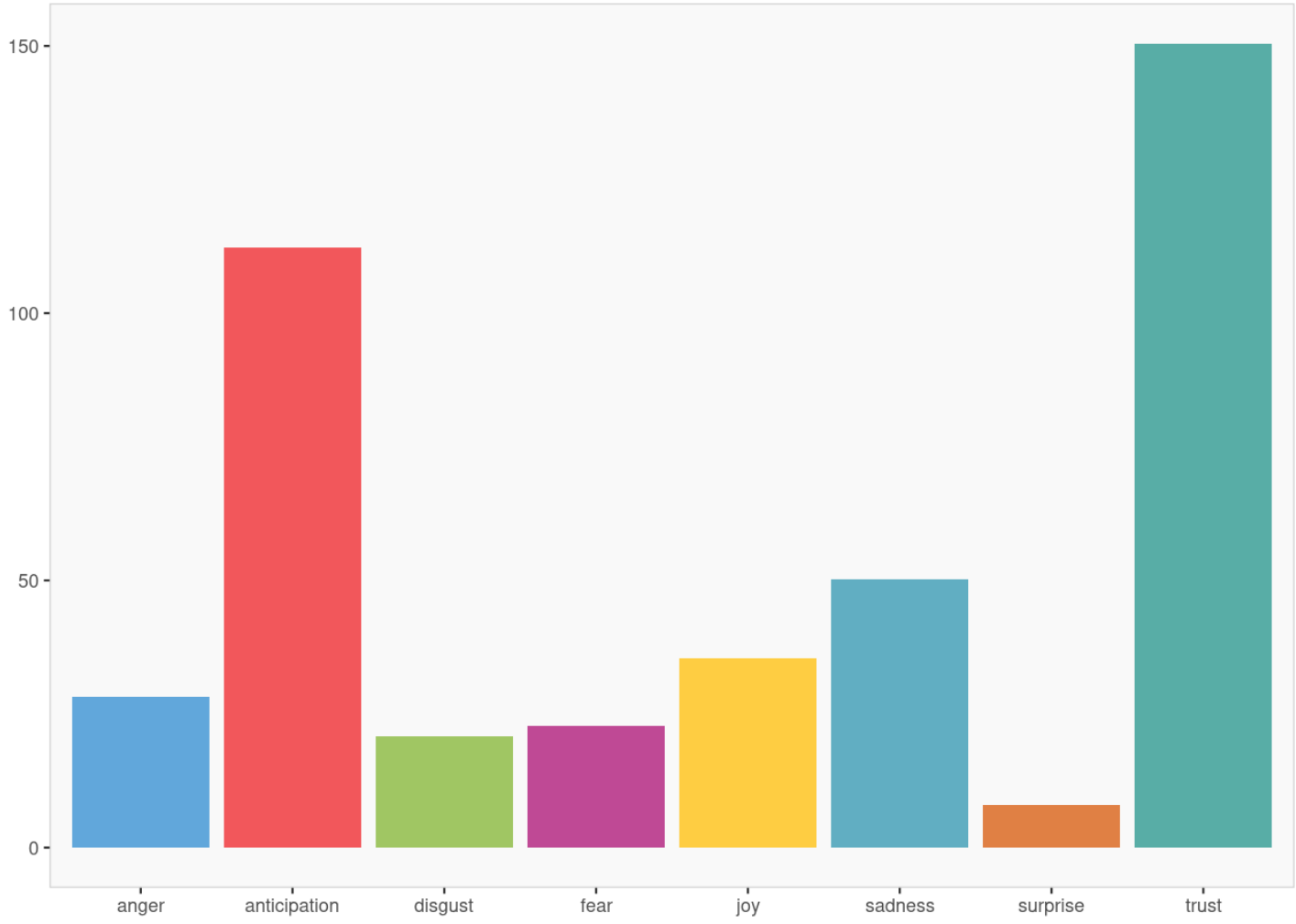

negative
neutral
positive
return
order
item
trouser
recommend
year
track
mask
face
jean
june
wait
bought
size
month
resel
august
shippingnow
today
septemb
paid
write
mind
chase
everlan
juli suitabl
futur
coat
poor
small
purchas
respond
respons
cancel
send
deliv
base
buy
shoe
left
receiv
time
compani
servic
custom
contact
week
ship
great
cloth
price
review
good
deliveri
pair
high
final
reach
sale
style
check
support
basiclabel
effici
imposs
design
white
strang
dress
retail
larg
tshirt
surpris
quick
box
denim
experi
work
love
product
ethic
qualit
nice
happi
brand
help
free
super
amaz
offer
want
edit
place
stock
cut
correct
black
fast
beautif
fine
provid
true
hole
measur
relax
rate
print
transact
request
cotton
suppos
transpar
expect
coupl
leg
messag
refund
cost
silk
updat
thought
abl
onlin
cheap
excus
simpli
handl
link
peopl
clear
bank
ridicul
answer
australian
consum
appli
communic
shipment
card
pay
wrong
reason
form
told
text
warn
sign
shame
blame
email
call
delay
credit
bad
issu
lost
phone
mistak
scam
leather
leav
money
absolut
shop
charg
wear
longer
refus
expens
transit
info
detail
overpr
hold
happen
complet
total
fee
boot
avoid
disappoint
account
packag
arriv
terribl
horribl
repli
liter
site
multipl
discount
status
carrier
accept
promis
frustrat
nonexist
fashion
incorrect
friend
resolv
process
sweater
easi
tax
sell
lot
fit
shirt
turn
dhl
perfect
pretti
gift
short
cashmer
guess
improv
address
skinni
piec
paypal
parcel
list
excel
initi
replac
store
busi
feel
garment
post
dec
summer
includ
photo
option
state
negat
australia
human
wardrob
impress
start
wash
live
pill
bewar
hope
slow
button
materi
nope
bag
pant
read
hour
number
delay
confirm
deal
worst
long
problem
point
sustain
glitch
feb
stay
realiz
polici
horrend
wast
covid
exchang
care
sad
blazer
rain
heard
mail
team
jacket
stuck
discov
buyer
claim
ask
worth
end
clean
cleaning
cleaner
cleanse
cleansing
cleanser
cleanse
cleansing
cleanser
cleanse
cleansing
cleanser

Emotion analysis

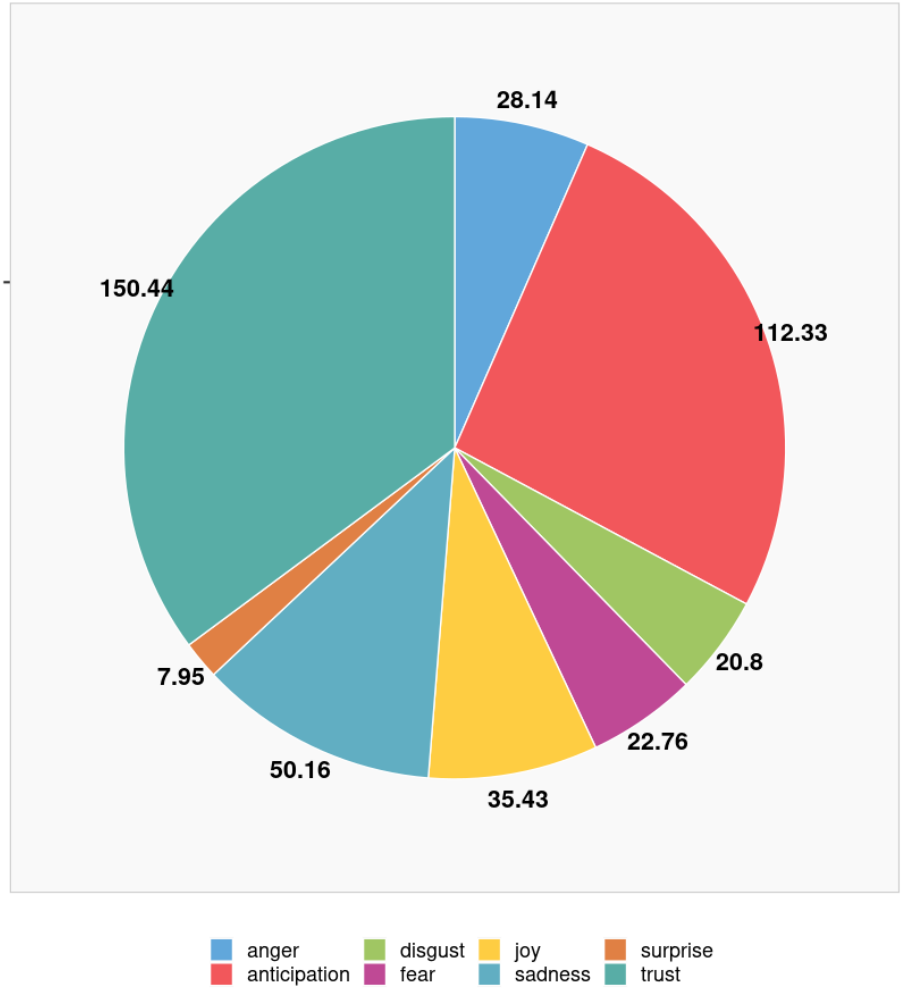
Emotion repartition

	Posts count	Relative posts count
Total	428	100%
anger	28.14	7%
anticipation	112.33	26%
disgust	20.8	5%
fear	22.76	5%
joy	35.43	8%
sadness	50.16	12%
surprise	7.95	2%
trust	150.44	35%

Emotion repartition. The number of posts that fall into different emotion categories summarized by their absolute and relative values. If a post has multiple emotions, then it is equally divided among those emotions.



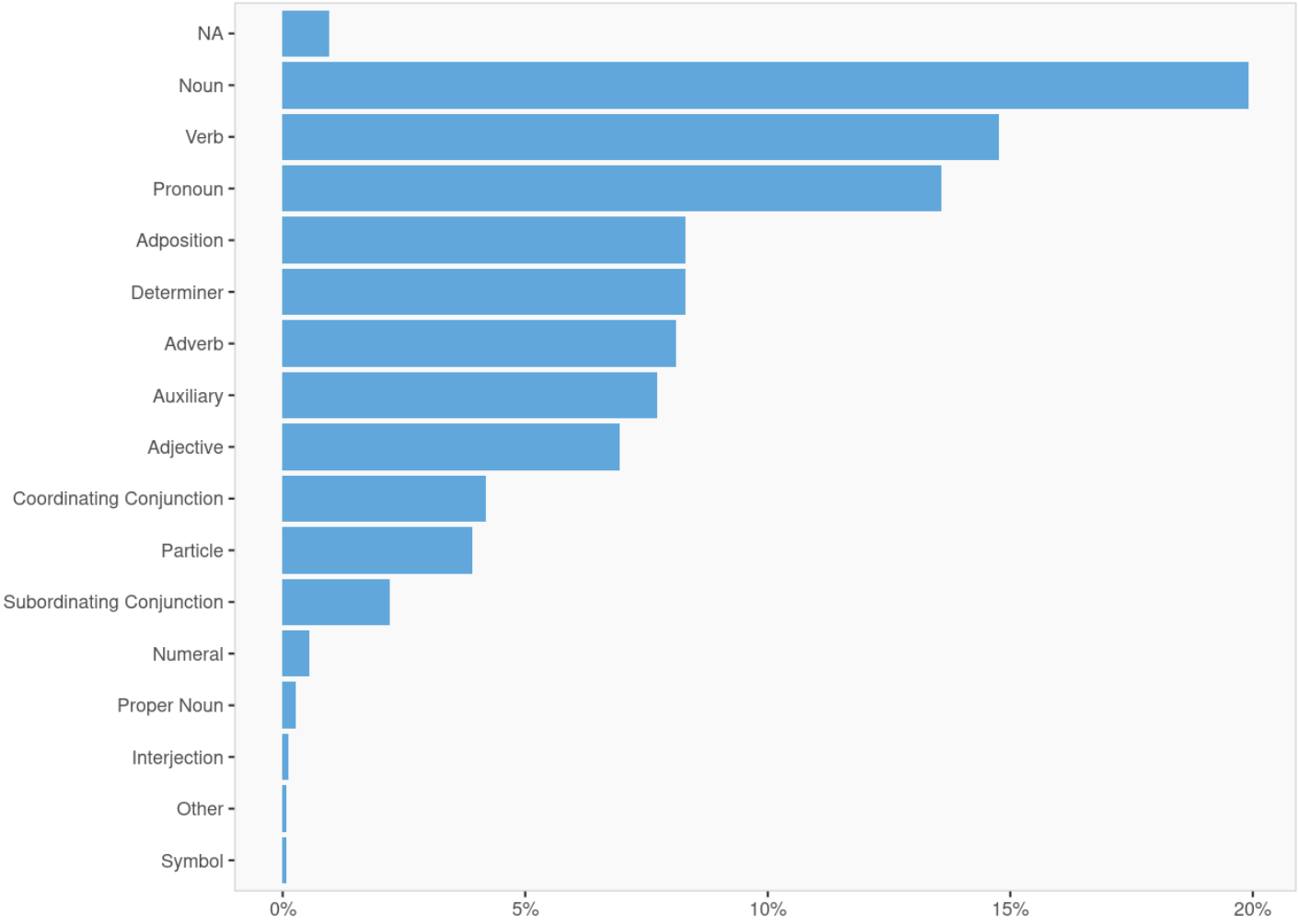
Emotion histogram. The emotion histogram indicates the number of posts by their emotion.



Emotion distribution without uncategorized posts. Graphic summary of the relative sizes of the number of posts classified by emotion after ignoring uncategorized posts.

Emotion word cloud

disgust
anticipation
fear
anger
joy
trust
sadness
surprise
great
avoid
qualiti
love
terribl
order
product
time
receivemail
servic
purchas
disappoint
everlan
charg
custom
credit
return
shirt
ship
long
week
shoe
boot
cost
absolut
item
jean
good
basic
tax
sale
price
process
excel
cashmer
sweater
denim
repli
cloth
dec
wash
sandal
shame
problem
deliveri
trash
read
develop
accept
complet
want
horribl
compani
slipper
horrend
ridicul
buyer
carrier
mail
discount
cancel
address
label
brand
card
respons
refund
recommend
communic
scam
guess
nice
xmas
size
account
packag
feel
care
final
cut
mouth
pant
arriv
excus
tshirt
pay
look
stock
reason
fur
step
everlast
blame
bot
blah
glitch
fraud
poor
short
follow
today
fee
parcel
wait
pretti
track
check
thought
buy
sustain
face
entir
jacket
unbeliev
day
adv
amex
elig
tear
construct
adjust
nervous
goodcustom
wors
happen
fals
stick
fact
huge
expens
slow
mediocr
difficult
wast
poshmark
union
rubbish
dread
employe
modern
yeah
switzerland
stitch
profit
enter
bought
doorstep
excit
resel
stay
exchang
fraudul
deserv
boss
restock
theft
stain
polish
answer
add
reach
mule
unauthor
dishonest
singl
sign
late
hole
delay
site
tee
detail
overpr
easili
link
market
wore
joy
question
appal
earli
thiev
worth
updat
sell
post
industri
wear
resolut
spar
chino
happi
pair
transp
help
easi
piec
free
busi
coupl
fan
leav
turn
bag
issu
account
black
colour
eye
star
big
red
fit
bad
worn
lost
deliv
garment
beauti
wardrob
review
high
super
friend
clear
outfit
comfort
model
courier
photo
dhl
actual
polici
top
notific
awesom
log
marin
experi
worst
strang
larg
decent
june
quick
surpris
sad
support
send
slip
tote
tank
fast
offer
medium
extra
drop
amaz
live
shop
nonexist
impress
negat
contact
multipl
shapeless
heel
base
classic
retail
favorit
year
money
idea
ethic
dress
review
wardrobe
clothing
shoes
jewelry
accessories
fashion
quality
price
product
shipping
order
delivery
customer
service
refund
return
exchange
scam
fraud
fake
cheap
expensive
worth
money
waste
broken
damaged
defective
disappointed
disappointing
terrible
horrible
awful
bad
good
great
excellent
perfect
love
hate
like
dislike
recommend
buy
purchase
order
receive
ship
delivery
late
slow
fast
quick
price



Distribution of Universal Parts of Speech Tags.

Top words

	NOUN	VERB	PRON	ADP	DET
1	customer	ordered	i	of	the
2	order	have	it	in	a
3	service	had	my	for	no
4	quality	received	they	from	an
5	company	get	them	on	this
6	shipping	return	their	to	all
7	items	took	everlane	with	some
8	time	got	me	at	any
9	item	order	you	out	another
10	refund	sent	that	about	these

Top words (excerpt).

Copyright (c) 2025, DecisionPro Inc.

enginius
MARKETING ENGINEERING ONLINE

Enginius

Sentiment Analysis

Charudisha Ashjay, University of Southern California

Copyright (c) 2025, DecisionPro Inc.

Sentiment options

Options selected

	Option	Selection
	Data source	Sentiment data
	Verbatim	text
	Word co-occurrence analysis and RAKE	No
	Topic model	No
	Default stop words	Long list (851)
	Custom stop words	No
	Date and time	2025-10-21 01:50:19 UTC

Options selected.

Data description

	Data	Number of Rows	Number of columns	Column names
1	Sentiment data	132	6	source_url, text, is_english_guess, score_numeric, date_parsed, ...

Data description.

Word cloud

Word cloud of the most frequent words

Word cloud. The word cloud represents the most frequently used words inside the corpus of texts provided. The bigger a word appears, the larger the number of times it occurs in the text corpus.

Word cloud of the most frequent words (without stemming)

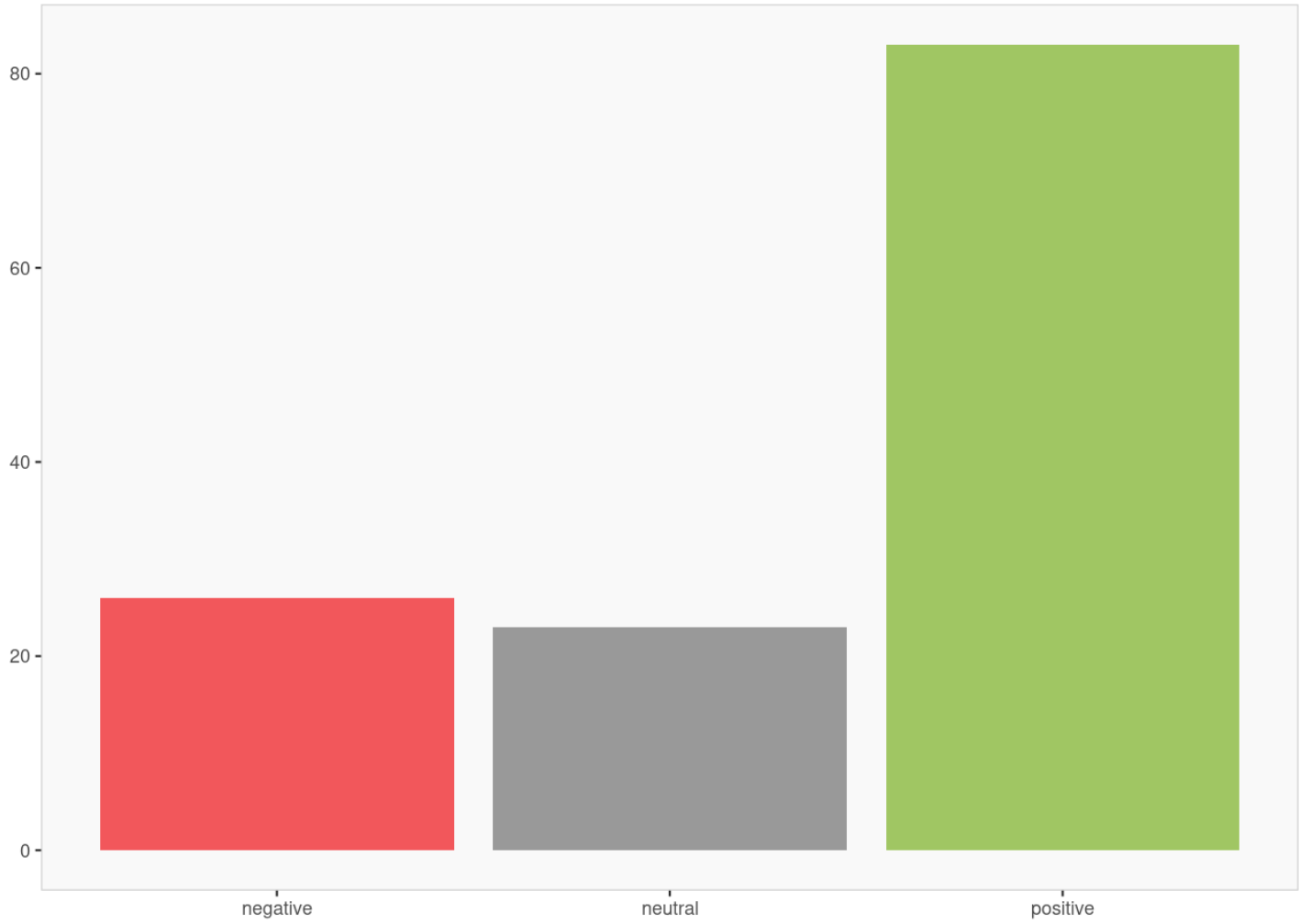
Word cloud without stemming. The word cloud represents the most frequently used words inside the corpus of texts provided. The bigger a word appears, the larger the number of times it occurs in the text corpus.

Valence analysis

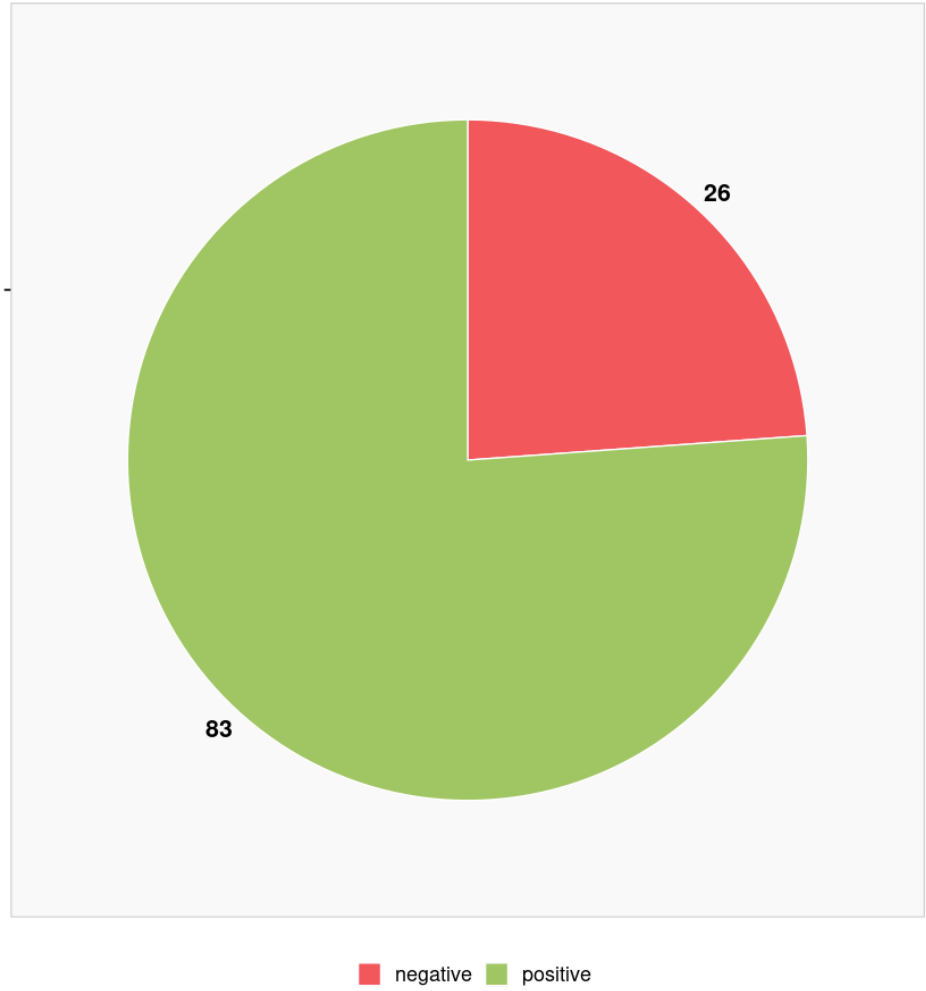
Valence repartition

	Posts count	Relative posts count
Total	132	100%
negative	26	20%
neutral	23	17%
positive	83	63%

Valence repartition. The number of posts that fall into different valence categories summarized by their absolute and relative values.

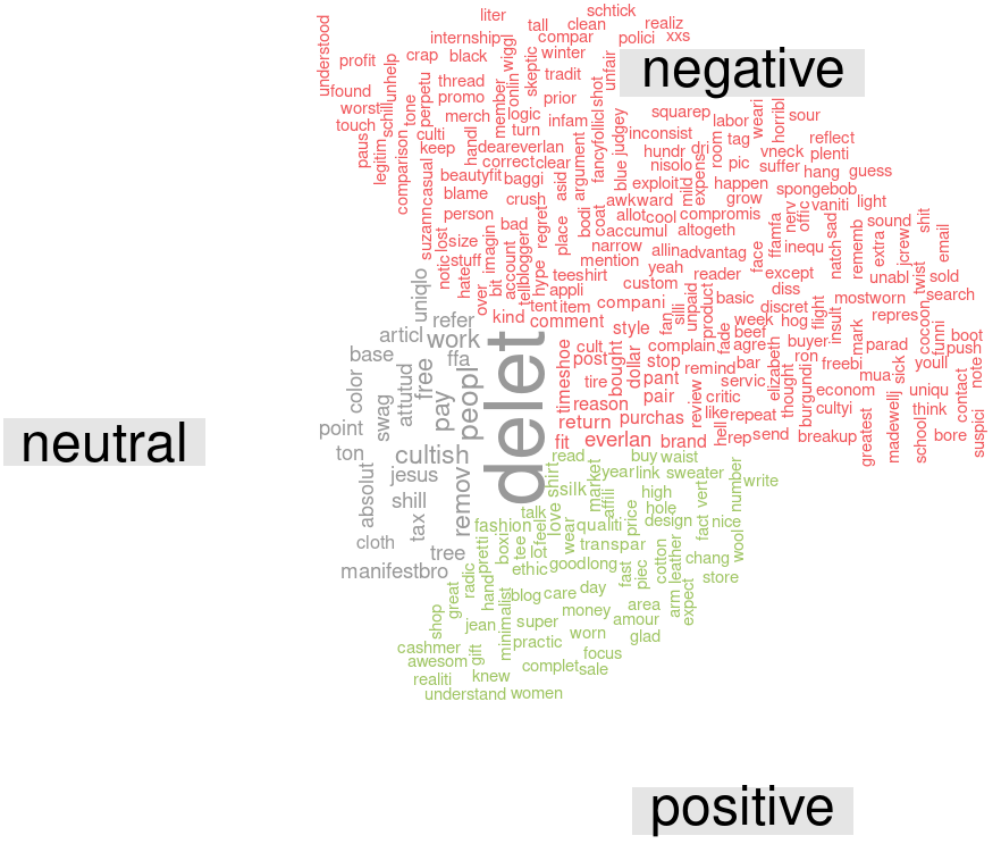


Valence histogram. The valence histogram indicates the number of posts by their valence.



Valence distribution without uncategorized posts. Graphic summary of the relative sizes of the number of posts classified by valence after ignoring posts that could not be categorized (i.e., neutral posts).

Valence word cloud



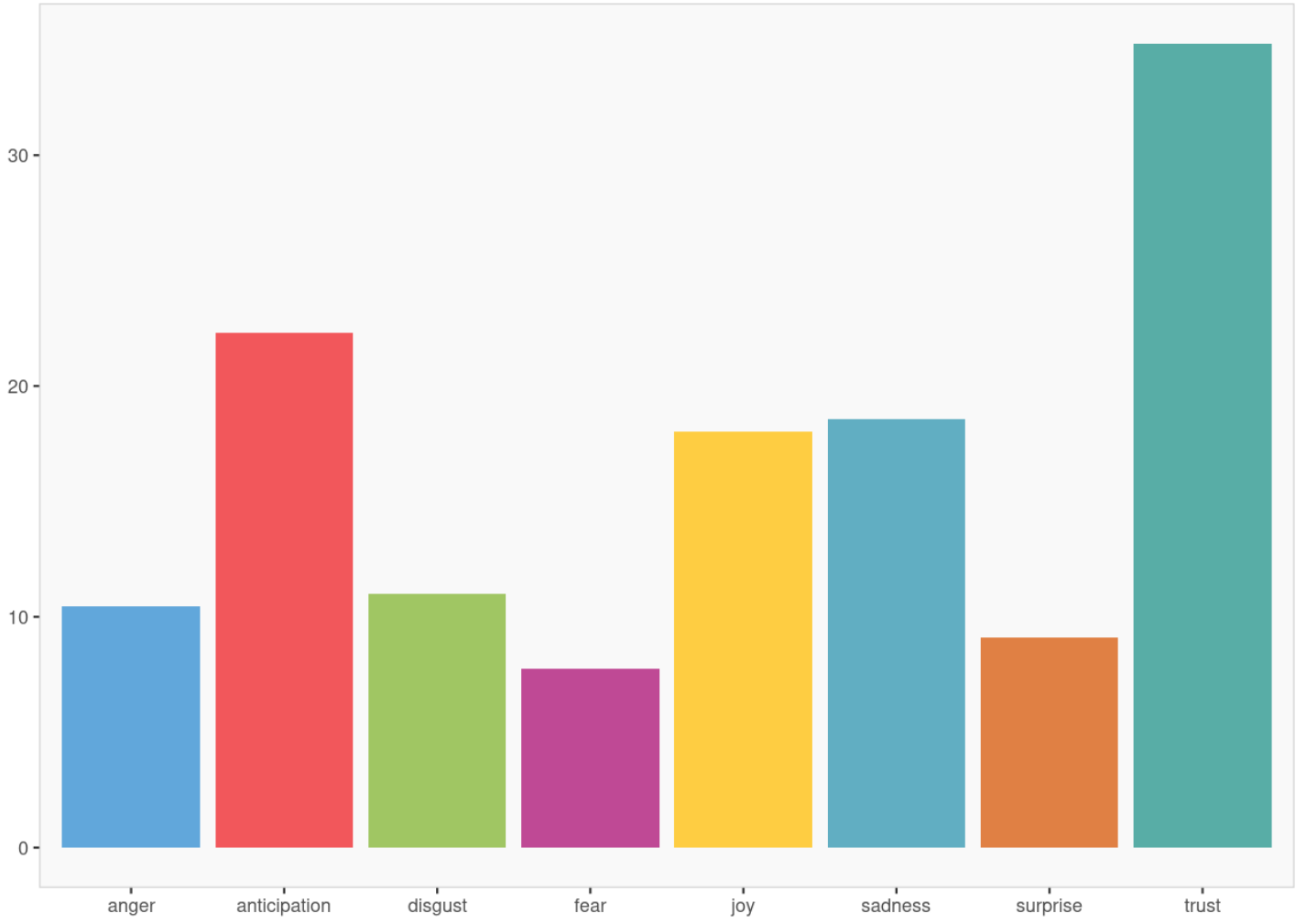
Valence word cloud.

Emotion analysis

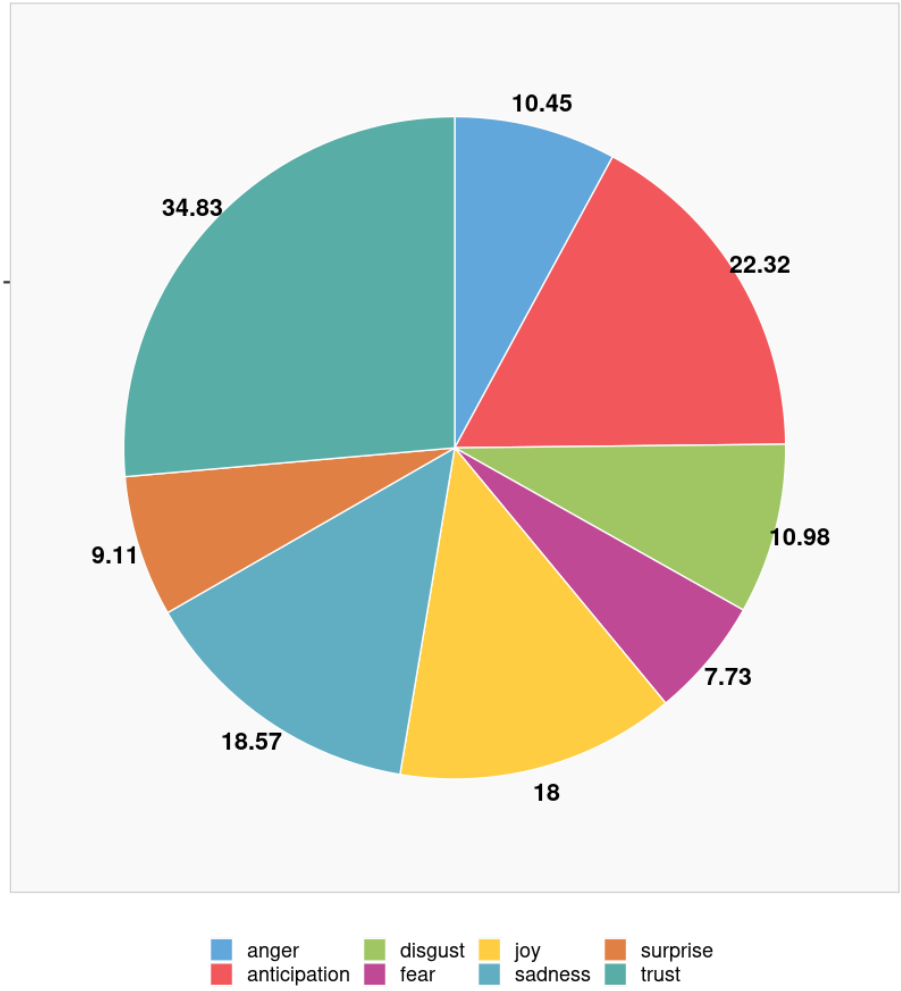
Emotion repartition

	Posts count	Relative posts count
Total	132	100%
anger	10.45	8%
anticipation	22.32	17%
disgust	10.98	8%
fear	7.73	6%
joy	18	14%
sadness	18.57	14%
surprise	9.11	7%
trust	34.83	26%

Emotion repartition. The number of posts that fall into different emotion categories summarized by their absolute and relative values. If a post has multiple emotions, then it is equally divided among those emotions.



Emotion histogram. The emotion histogram indicates the number of posts by their emotion.

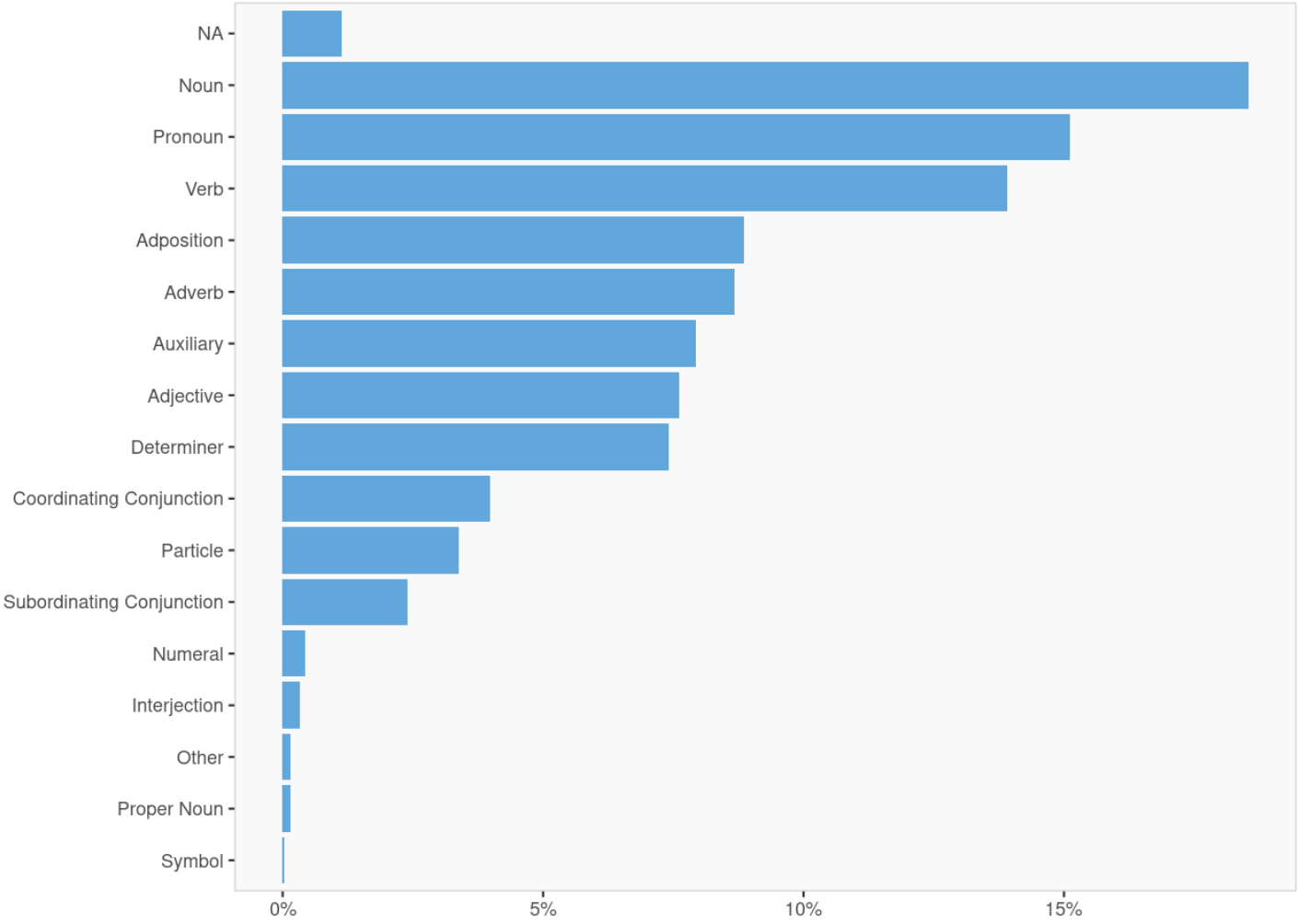


Emotion distribution without uncategorized posts. Graphic summary of the relative sizes of the number of posts classified by emotion after ignoring uncategorized posts.

Emotion word cloud

Emotion word cloud. Even if a post has multiple emotions its words will be shown only in one of those emotions.

Distribution of Universal Parts of Speech Tags



Distribution of Universal Parts of Speech Tags.

Top words

	NOUN	PRON	VERB	ADP	ADV
1	quality	i	have	of	just
2	clothes	it	think	for	really
3	fashion	they	like	in	so
4	brand	their	get	with	too
5	items	that	bought	from	more
6	everlane	you	make	on	very
7	people	them	know	about	when
8	brands	me	fit	at	why
9	lot	my	love	to	also
10	silk	she	deleted	like	still

Top words (excerpt).

Copyright (c) 2025, DecisionPro Inc.

Enginius

Segmentation

Charudisha Ashjay, University of Southern California

Copyright (c) 2025, DecisionPro Inc.

Segmentation options

Options selected

	Option	Selection
	Clustering method	Hierarchical
	Standardization method	none
	Segments forced	No
	Run discriminant analysis	Yes
	Run classification analysis	Yes
	Date and time	2025-12-03 06:52:42 UTC

Options selected.

Data description

	Data	Number of Rows	Number of columns	Column names
1	Segmentation data	68	11	RespID, Style, Quality, Fit, Value for Price, ...
2	Discriminant data	68	28	RespID, 1. Are you familiar with the brand Everlane?, 2. Have you ever purchased from Everlane?, 3. How aware are you of Everlane's sustainability or transparency initiatives (e.g., recycled materials, ethical factories)?, 4. In the past, how much have Everlane's sustainability claims influenced your likelihood of purchasing from them?, ...

Data description.

Data transformation

Standardization has not been performed.

Segment solution

3-segment solution

The ideal number of segments is a function of statistical fit (what the data say), managerial relevance (what makes the most sense from a managerial point of view), and targetability (can the segments be easily targeted).

When the three criteria do not perfectly converge, selecting the right number of segments becomes a judgment call.

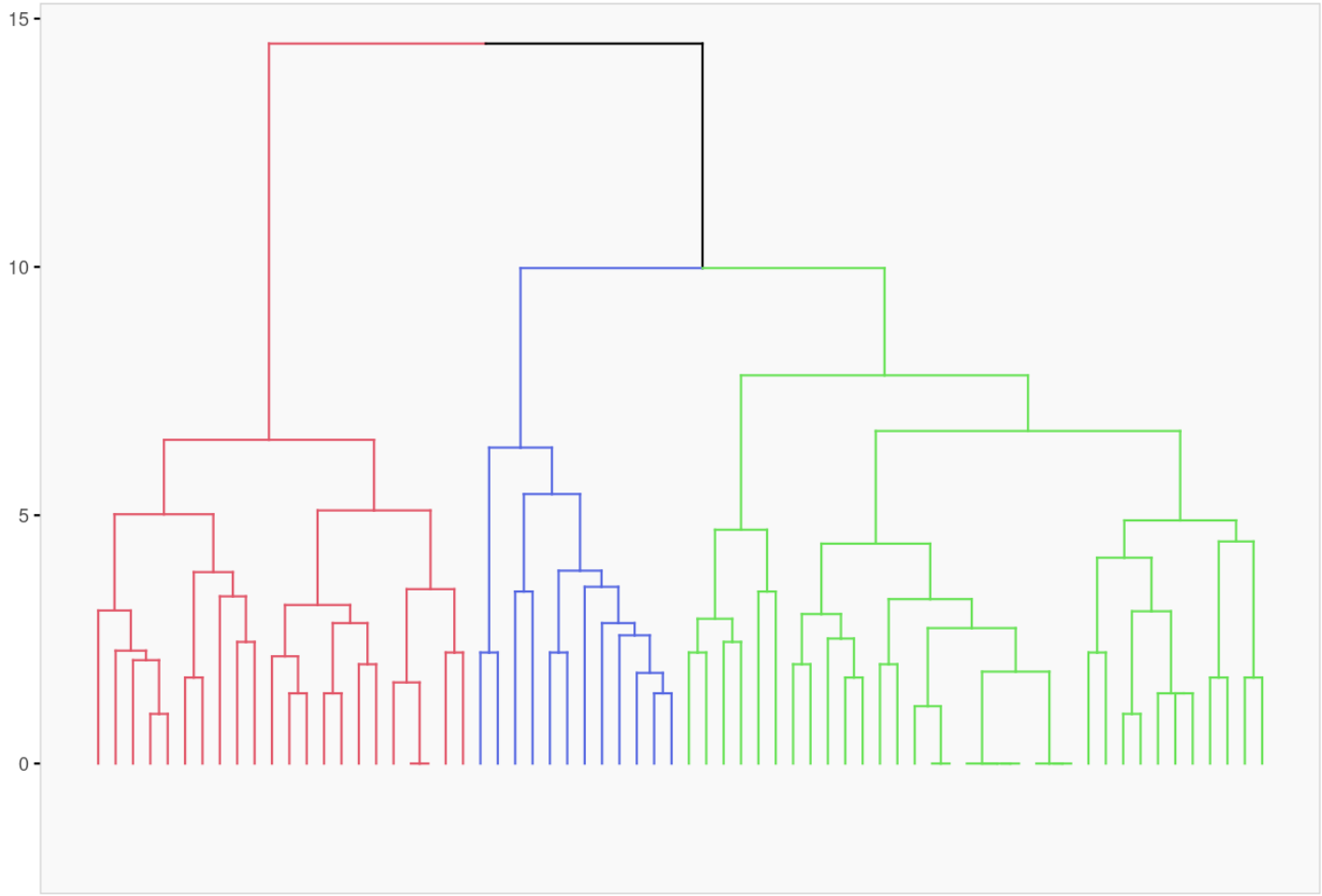
Using a statistical criteria exclusively (see scree plot analysis below), we have retained 3 segments.

The segmentation method relies on the hierarchical clustering approach. This approach generates a dendrogram that we display next.

Dendrogram

The dendrogram represents the grouping process of observations into clusters. The chart reads from bottom (all initial observations are separated) to top (all observations are clustered into one unique segment).

The height represents the distance between the two groups of observations being merged at each step. If two very distant groups are being merged, this will create a 'jump' in the dendrogram, indicating that it might be wise to stop the clustering process before.



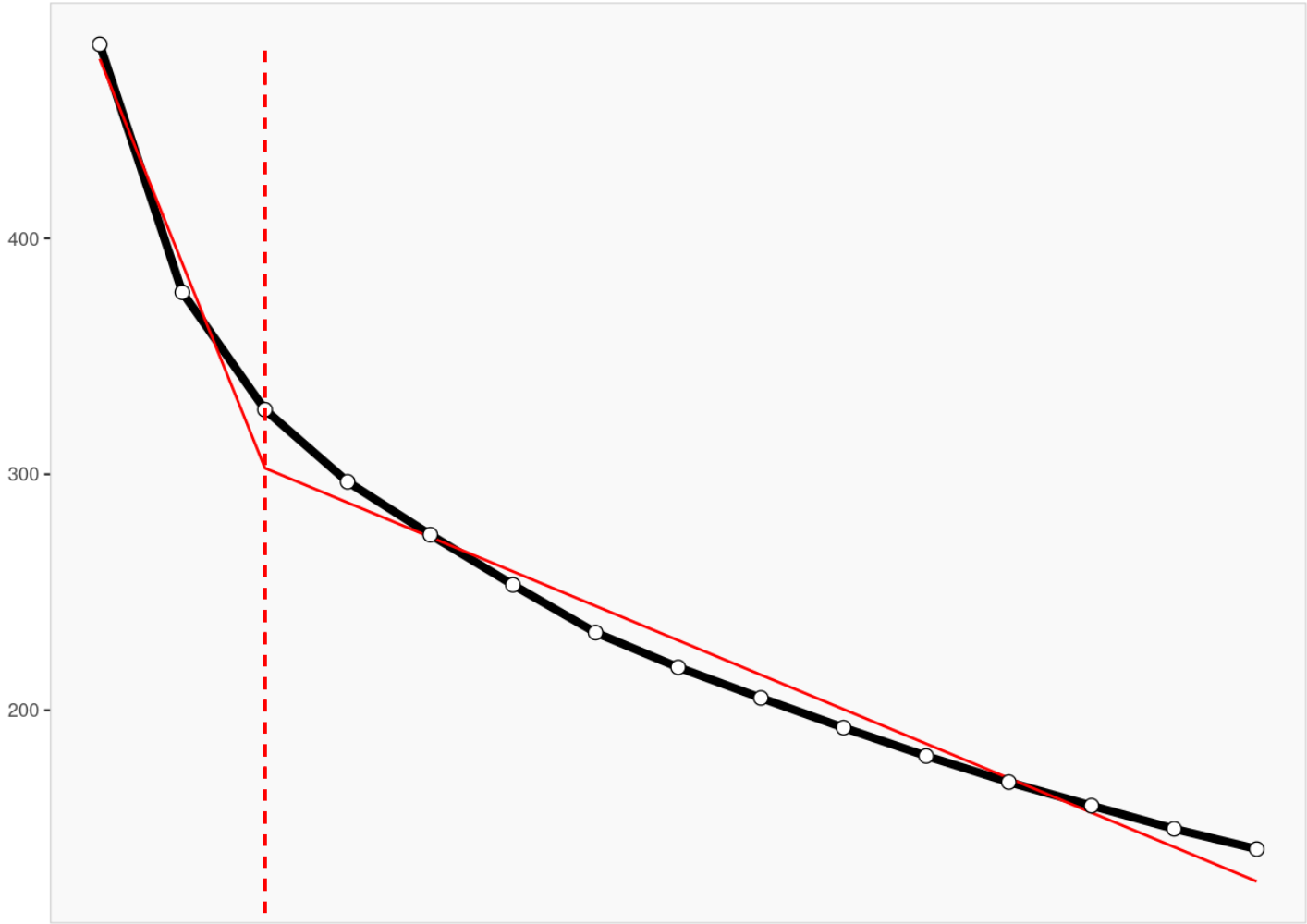
Dendrogram. The dendrogram is a tree diagram to illustrate the arrangement of clusters produced by hierarchical clustering, and how the observations are incrementally clustered together.

Scree plot

The screeplot displays, for each cluster solution, a measure of within-cluster heterogeneity. If clusters group observations that are widely different (which will happen if the number of clusters is too small to capture the variability in the data), the value will be high.

A good cluster solution might be where the screeplot displays an 'elbow', that is, where increasing the number of clusters beyond a certain point does not dramatically decreases within-cluster heterogeneity.

The measure displayed in the screeplot is related, but not equivalent, to the distance reported in the dendrogram.



Scree plot. The scree plot compares the sum of squared error (SSE) for each cluster solution. A good cluster solution might be when the SSE slows dramatically, creating an 'elbow'. Such elbow does not always exist. If number of segments is equal to maxumum possible segments elbow cannot be created.

From a statistical point of view, the SSE reported in the screeplot is computed as the sum of squared error between each observation and its cluster centroid (or center), summed over all the observations.

Segment description

Segment size

	Population	Segment 1	Segment 2	Segment 3
Size	68	22	12	34
Relative size	100%	32%	18%	50%

Segment size.

Segment description

	Population	Segment 1	Segment 2	Segment 3
Style	4.4	4.45	4.08	4.47
Quality	4.43	4.77	4.67	4.12
Fit	4.54	4.95	4.92	4.15
Value for Price	4.25	4.73	4	4.03
Durability	3.99	4.82	4.08	3.41
Sustainability	3.44	4.27	2.25	3.32
Trendy	3.69	4.05	2.42	3.91
Versatile	3.9	4.59	4.08	3.38
Brand reputation	3.32	4	2.83	3.06
Ease of shopping	3.9	4.41	3.33	3.76

Segment description. Average value of each segmentation variable, overall for each segment (centroid). Segmentation variables that are statistically different from the rest of the population are highlighted in red (lower) or green (higher).

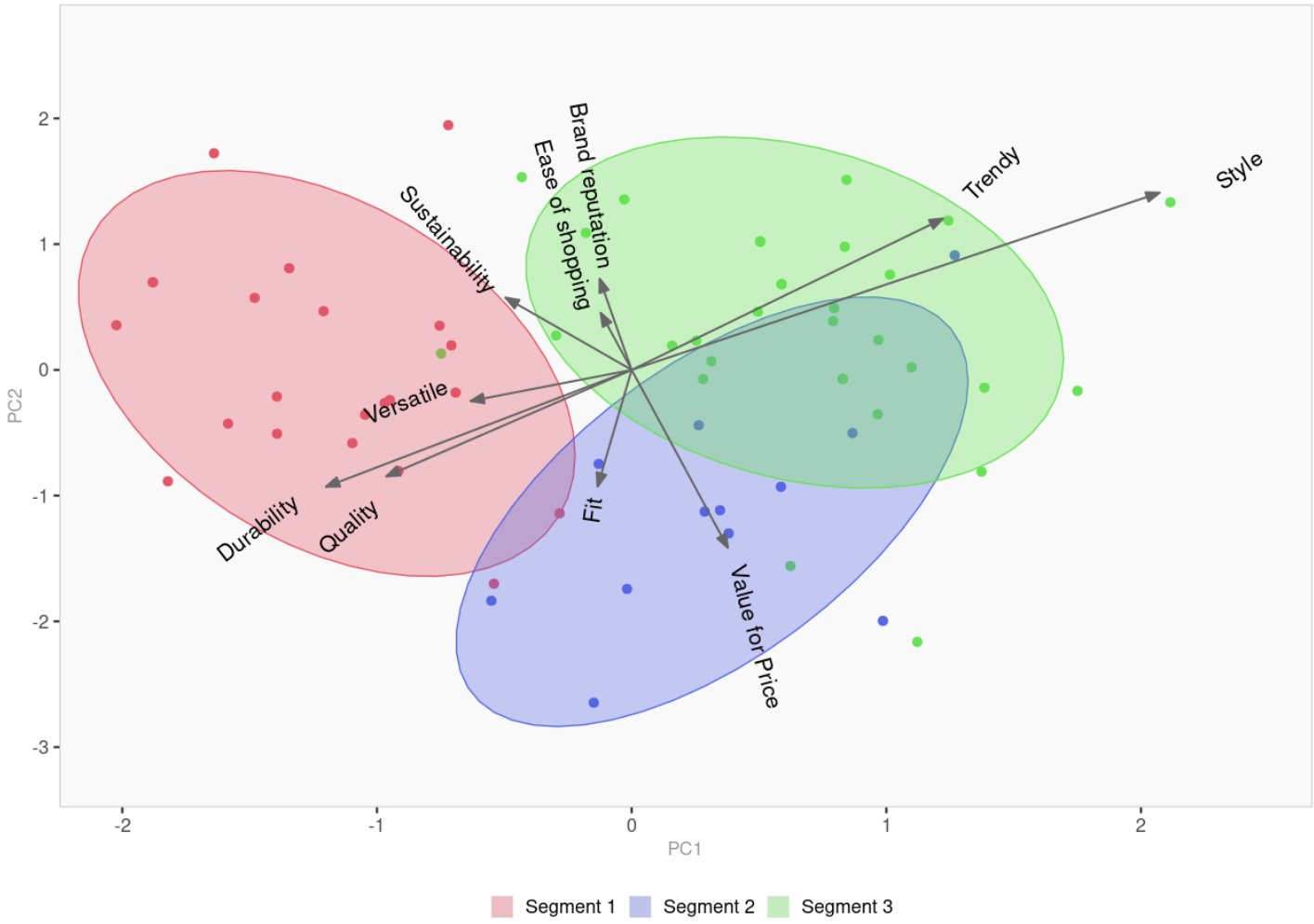
Segment differences per segment. Cell colors indicate to what extent a segment is statistically different from the rest of the population on each segmentation variable.

Segmentation space

The chart below is a graphical representation of the various segments, segment members, and segmentation variables. It is obtained by plotting the first two dimensions of a principal component analysis performed on the (standardized) segmentation data, on top of which segment information has been overlaid.

Because only the first two dimensions of the PCA are displayed, and these two dimensions capture only 47.3% of the variance in the data, some differences between segments might not appear here. Note that segmentation variables with no variance, if any, have been excluded.

Two clusters that appear to overlap on the first two dimensions might be distinct on other dimensions. Consequently, this chart is a useful guide, for checking which variables are correlated, but may be misleading if used to select the optimal number of segments.



Segment space. Spatial representation of segments and segmentation variables, using principal component analysis.

Segment membership

	Segment
1	3
2	3
3	2
4	3
5	2
6	1
7	1
8	3
9	1
10	3

Segment membership (excerpt). Segment to which each member of the population belongs to. The complete membership list is only available in the Excel formatted output.

Segment profiles

Spider chart

Spider chart comparing the averages of the segmentation variables across all segments.

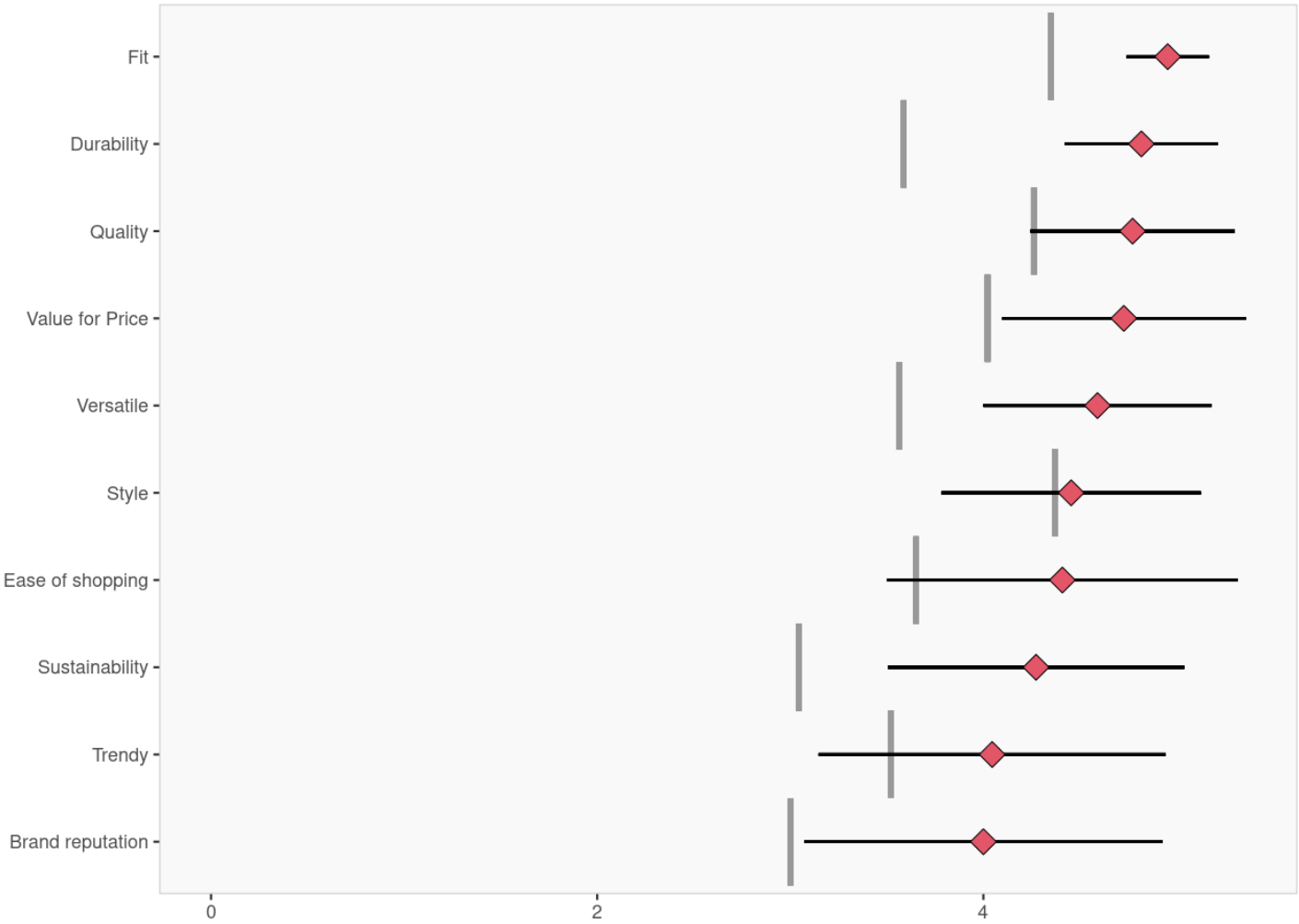
Spider chart.

Segment 1 profile

The following charts represent the profile of each segment. These charts are only available when the data are not standardized, hence the model assumes that all segmentation variables use the same scale.

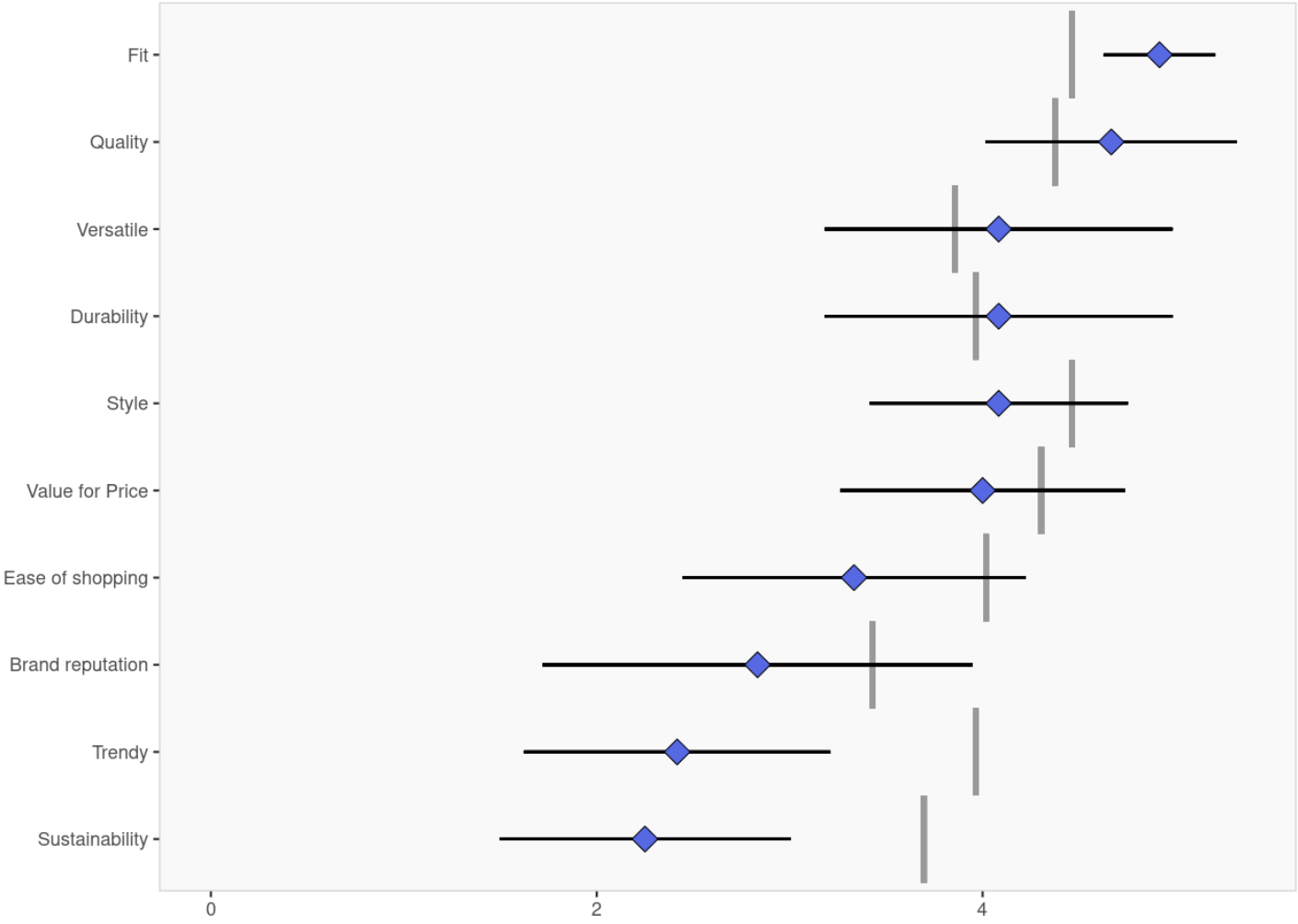
- For each segment, the segmentation variables are ordered in decreasing order of magnitude.
- The colored dots represent the average of the segment.
- The horizontal lines represent the standard deviations within that segment.
- The vertical, gray lines represent the averages of the rest of the population, after excluding members of

the segment under scrutiny.



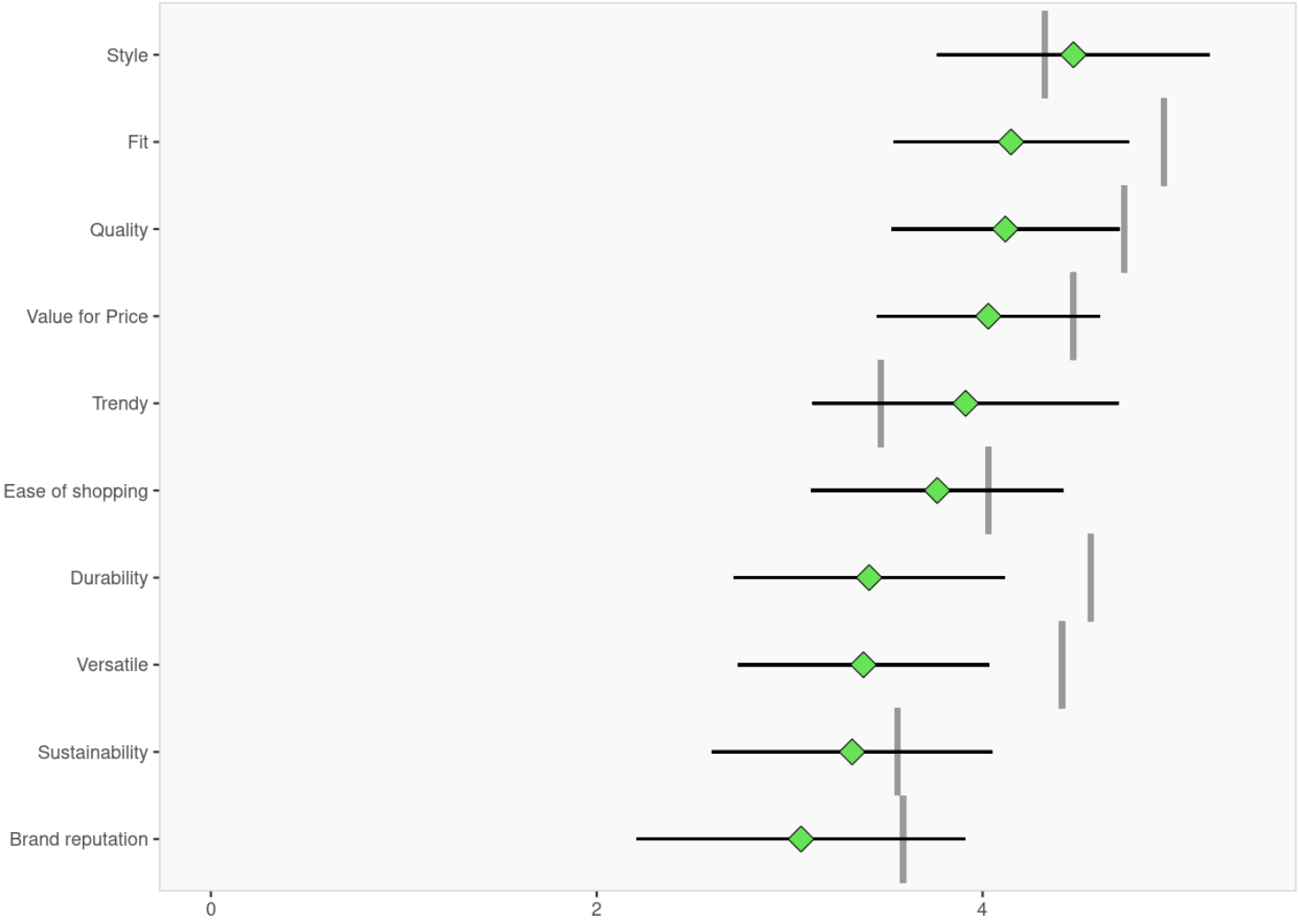
Segment 1 profile.

Segment 2 profile



Segment 2 profile.

Segment 3 profile



Segment 3 profile.

Descriptor analysis

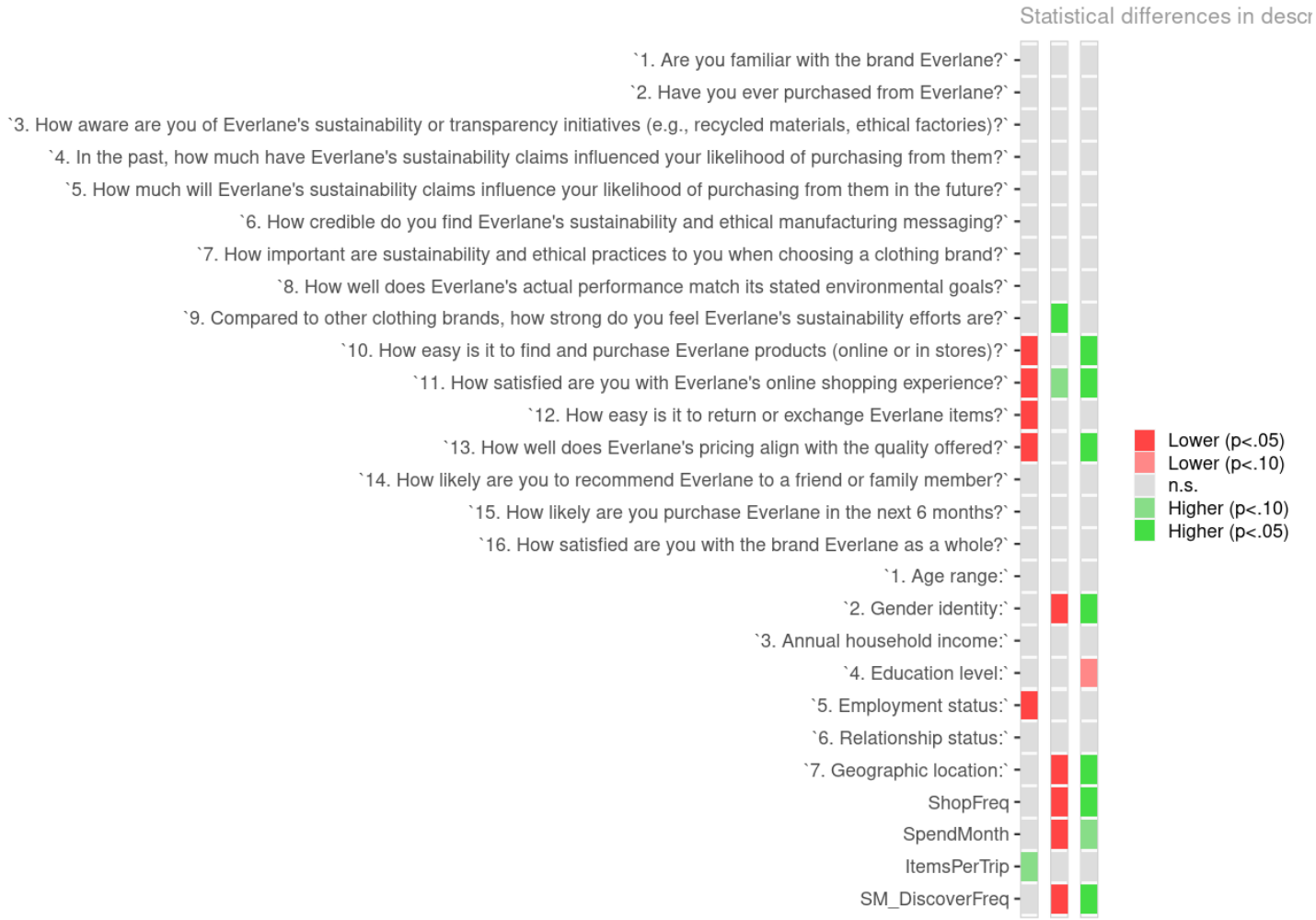
Descriptors

This table reports the descriptor averages of each segment. The more differences can be found, the easier it will be to predict segment membership based on descriptors alone.

	Population	Segment 1	Segment 2	Segment 3
`1. Are you familiar with the brand Everlane?`	0.353	0.227	0.333	0.441
`2. Have you ever purchased from Everlane?`	0.191	0.091	0.25	0.235
`3. How aware are you of Everlane's sustainability or transparency initiatives (e.g., recycled materials, ethical factories)?`	1.6	1.5	2.17	1.47
`4. In the past, how much have Everlane's sustainability claims influenced your likelihood of purchasing from them?`	1.79	1.64	1.83	1.88
`5. How much will Everlane's sustainability claims influence your likelihood of purchasing from them in the future?`	2.25	2.09	2.25	2.35
`6. How credible do you find Everlane's sustainability and ethical manufacturing messaging?`	2.46	2.32	2.25	2.62
`7. How important are sustainability and ethical practices to you when choosing a clothing brand?`	2.94	3.09	2.75	2.91
`8. How well does Everlane's actual performance match its stated environmental goals?`	2.65	2.68	2.5	2.68
`9. Compared to other clothing brands, how strong do you feel Everlane's sustainability efforts are?`	2.93	2.64	3.5	2.91
`10. How easy is it to find and purchase Everlane	3.07	2.68	3.17	3.29

products (online or in stores)?`				
`11. How satisfied are you with Everlane's online shopping experience?`	3.15	2.64	3.42	3.38
`12. How easy is it to return or exchange Everlane items?`	3.24	2.95	3.58	3.29
`13. How well does Everlane's pricing align with the quality offered?`	3	2.55	3	3.29
`14. How likely are you to recommend Everlane to a friend or family member?`	2.84	2.82	3.08	2.76
`15. How likely are you purchase Everlane in the next 6 months?`	2.63	2.5	2.58	2.74
`16. How satisfied are you with the brand Everlane as a whole?`	2.71	2.68	2.92	2.65
`1. Age range:`	1.59	1.5	1.75	1.59
`2. Gender identity:`	0.515	0.5	0.167	0.647
`3. Annual household income:`	2.35	2	2.83	2.41
`4. Education level:`	2.43	2.59	2.58	2.26
`5. Employment status:`	1.97	1.55	2.33	2.12
`6. Relationship status:`	1.49	1.36	1.75	1.47
`7. Geographic location:`	1.18	1.09	1	1.29
ShopFreq	3.49	3.23	2.58	3.97
SpendMonth	2.16	2.18	1.5	2.38
ItemsPerTrip	1.75	2	1.42	1.71
SM_DiscoverFreq	3.51	3.45	2.33	3.97

Descriptor data per segment. Average value of each descriptor, overall and within each cluster. Descriptors that are statistically different from the rest of the population are highlighted in red (lower) or green (higher).



Descriptor differences per segment. Cell colors indicate to what extent the distribution of a descriptor in a segment is statistically different from the rest of the population.

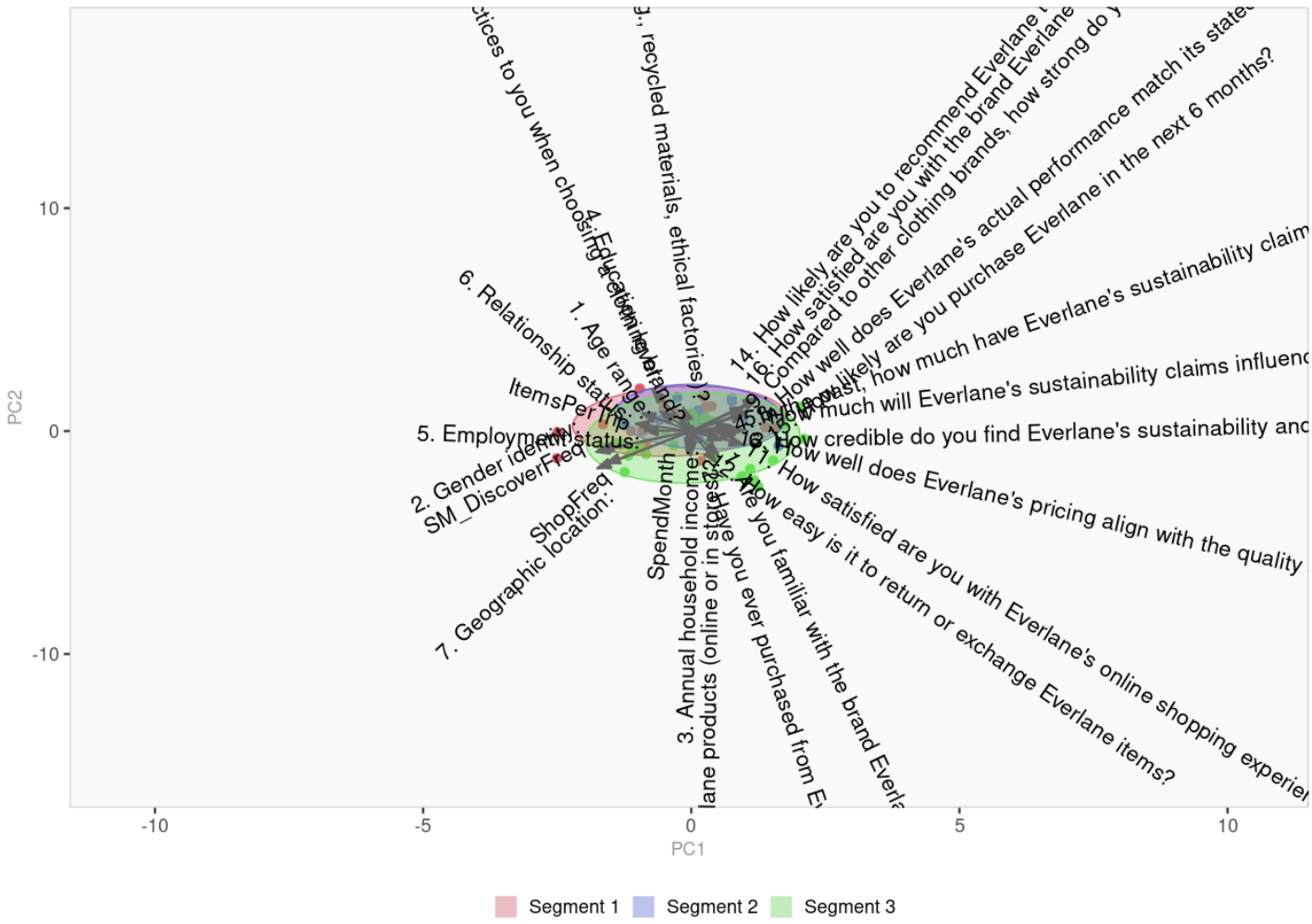
Descriptor space

The chart below is a graphical representation of the various segments, segment members, and descriptors. It is obtained by outputing the first two dimensions of a principal component analysis performed on the (standardized) descriptors, on top of which segment information has been overlayed.

Because only the first two dimensions of the PCA are displayed, and these two dimensions capture only 36.5% of the variance in the data, some differences between segments might not appear here. Note that descriptors with no variance, if any, have been excluded.

If two or more segments fully overlap, it is unlikely that they could be clearly separated based on descriptors alone.

However, two segments that seem to overlap on two dimensions may be more clearly separated on other dimensions. Consequently, the confusion matrix is a better guide to assess the quality of segment classification based on descriptors.



Descriptor space. Spatial representation of segments and their descriptors, using principal component analysis.

Classification model

Introduction

Often, segmentation (needs) variables for each customer may not be available to managers, but descriptors variables for customers may be available.

In this section, we explore whether descriptors alone can predict segment membership with sufficient accuracy. The confusion matrix and hit rates (reported below) indicate whether the model is accurate enough.

For member classification based on descriptors, Enginius uses a multinomial logit model (similar to the one used to predict 'choices between multiple alternatives (A/B/C)' in the predictive modeling module.

The largest segment is selected as the default option (dummy), and the model identifies which descriptors are the most significant for predicting cluster memberships. If a descriptor is highly predictive, its p-values will be close to zero, and the cells will appear in green (or red).

Model coefficients

	Segment 1	Segment 2
(Intercept)	2 638	168
1. Are you familiar with the brand Everlane?	-509	-1 233
2. Have you ever purchased from Everlane?	519	127
3. How aware are you of Everlane's sustainability or transparency initiatives (e.g., recycled materials, ethical factories)?	-358	-128
4. In the past, how much have Everlane's sustainability claims influenced your likelihood of purchasing from them?	120.2	68.9
5. How much will Everlane's sustainability claims influence your likelihood of purchasing from them in the future?	158	742
6. How credible do you find Everlane's sustainability and ethical manufacturing messaging?	-86.5	62.9
7. How important are sustainability and ethical practices to you when choosing a clothing brand?	77.5	148.1
8. How well does Everlane's actual performance match its stated environmental goals?	-139	-890
9. Compared to other clothing brands, how strong do you feel Everlane's sustainability efforts are?	-59.4	546.8
10. How easy is it to find and purchase Everlane products (online or in stores)?	-10.3	-447.5
11. How satisfied are you with Everlane's online shopping experience?	-318	898

12. How easy is it to return or exchange Everlane items?	132	661
13. How well does Everlane's pricing align with the quality offered?	-754	-1 006
14. How likely are you to recommend Everlane to a friend or family member?	908	271
15. How likely are you purchase Everlane in the next 6 months?	-246	-449
16. How satisfied are you with the brand Everlane as a whole?	98.7	387.6
1. Age range:	-155.74	2.25
2. Gender identity:	-103	-959
3. Annual household income:	125.6	86.3
4. Education level:	44.5	-460.9
5. Employment status:	-169.5	-57.9
6. Relationship status:	-360	178
7. Geographic location:	92.1	-251.5
ShopFreq	-160	-502
SpendMonth	-27.7	-375.6
ItemsPerTrip	255	412
SM_DiscoverFreq	-186.8	49.8

Model parameters. Segment 3 is the model baseline.

P-values

	Segment 1	Segment 2
(Intercept)	0	0
1. Are you familiar with the brand Everlane?	0	0
2. Have you ever purchased from Everlane?	0	0
3. How aware are you of Everlane's sustainability or transparency initiatives (e.g., recycled materials, ethical factories)?	0.971	0
4. In the past, how much have Everlane's sustainability claims influenced your likelihood of purchasing from them?	0.983	0
5. How much will Everlane's sustainability claims influence your likelihood of purchasing from them in the future?	0.932	0
6. How credible do you find Everlane's sustainability and ethical manufacturing messaging?	0.939	0
7. How important are sustainability and ethical practices to you when choosing a clothing brand?	0.969	0
8. How well does Everlane's actual performance match its stated environmental goals?	0.988	0
9. Compared to other clothing brands, how strong do you feel Everlane's sustainability efforts are?	0.977	0

10. How easy is it to find and purchase Everlane products (online or in stores)?	0.995	0
11. How satisfied are you with Everlane's online shopping experience?	0.874	0
12. How easy is it to return or exchange Everlane items?	0.918	0
13. How well does Everlane's pricing align with the quality offered?	0.9	0
14. How likely are you to recommend Everlane to a friend or family member?	0.647	0
15. How likely are you purchase Everlane in the next 6 months?	0.897	0
16. How satisfied are you with the brand Everlane as a whole?	0.987	0
1. Age range:	0.755	0
2. Gender identity:	0.242	0
3. Annual household income:	0.988	0
4. Education level:	0.987	0
5. Employment status:	0.976	0
6. Relationship status:	0.536	0
7. Geographic location:	0.991	0
ShopFreq	0.976	0
SpendMonth	0.953	0
ItemsPerTrip	0.603	0
SM_DiscoverFreq	0.933	0

p-values. Probabilities that parameter estimates are different from zero only by chance.

Confusion matrix

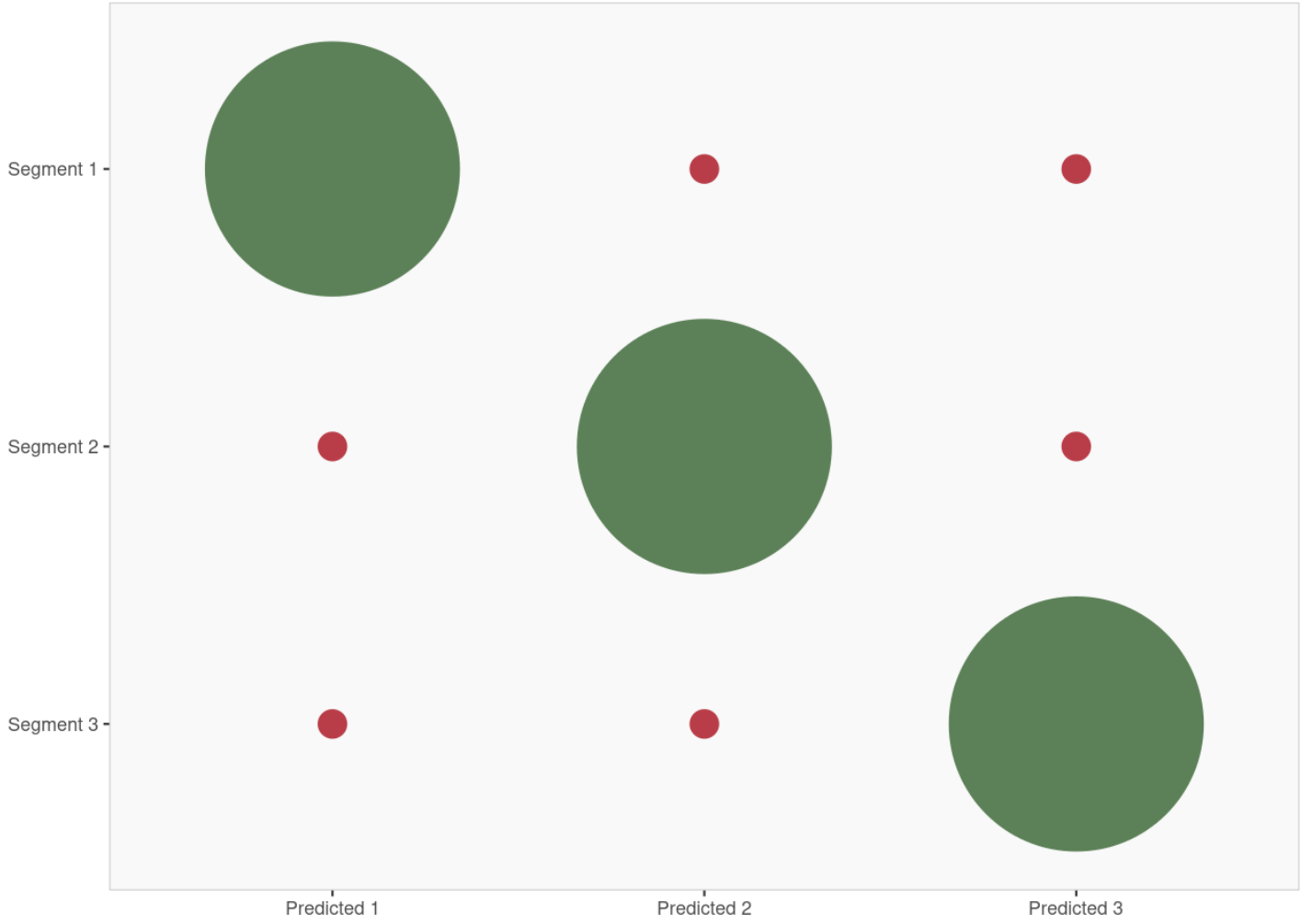
The confusion matrix compares actual segment membership (obtained from the segmentation analysis and the original segmentation variables) and predicted segment membership (obtained from the in-sample classification analysis and the descriptors alone). When actual and predicted segment memberships coincide, the diagonal elements will be comparatively large, indicating that the classification model based on available descriptors is accurate.

	Predicted 1	Predicted 2	Predicted 3	Total
Segment 1	22	0	0	22
Segment 2	0	12	0	12
Segment 3	0	0	34	34
Total	22	12	34	68

Confusion matrix (count). The model has correctly classified 68 of the 68 observations. The off-diagonal elements are classification errors.

	Predicted 1	Predicted 2	Predicted 3	Total
Segment 1	100%	0%	0%	100%
Segment 2	0%	100%	0%	100%
Segment 3	0%	0%	100%	100%

Confusion matrix (%). The global hit rate of the model is 100%. The diagonal elements represent segment-specific hit rates.



Confusion matrix (plot). Graphic representation of the confusion matrix: actual segment membership versus predicted segment membership. Bubbles in the diagonale represent correct classification.

Model predictions

	Prob(cluster 1)	Prob(cluster 2)	Prob(cluster 3)	Predicted	Actual	Correct
1	0%	0%	100%	3	3	1
2	0%	0%	100%	3	3	1
3	0%	100%	0%	2	2	1
4	0%	0%	100%	3	3	1
5	0%	100%	0%	2	2	1
6	100%	0%	0%	1	1	1
7	100%	0%	0%	1	1	1
8	0%	0%	100%	3	3	1
9	100%	0%	0%	1	1	1
10	0%	0%	100%	3	3	1

Model predictions (in-sample) (excerpt). This table details the probabilities of each member of the segmentation dataset to belong to each cluster (as predicted by the in-sample classification model and the descriptors alone). The segment with the highest probability is retained, and is compared to the actual segment membership to measure model accuracy and classification errors.

Copyright (c) 2025, DecisionPro Inc.

Factor Analysis

Eigenvalues

Number	Eigenvalue	Percent	20 40 60 80	Cum Percent
1	4.8655	48.655		48.655
2	1.6970	16.970		65.626
3	0.8641	8.641		74.267
4	0.6297	6.297		80.564
5	0.4428	4.428		84.992
6	0.3952	3.952		88.944
7	0.3546	3.546		92.490
8	0.3209	3.209		95.698
9	0.2528	2.528		98.226
10	0.1774	1.774		100.000

Factor Analysis on Correlations with 3 Factors: Maximum Likelihood / Varim

Final Communality Estimates

1. Style.1	0.72650
2. Quality.1	0.80389
3. Fit on my Body.1	0.63930
4. Value for Price.1	0.53122
5. Durability.1	0.82342
6. Sustainability.1	0.56511
7. Trendy/fashionable.1	0.68346
8. Versatile/appropriate to wear daily.1	0.63436
9. Brand reputation.1	0.56673
10. Ease of shopping (online/in-store experience).1	0.40715

Variance Explained by Each Factor

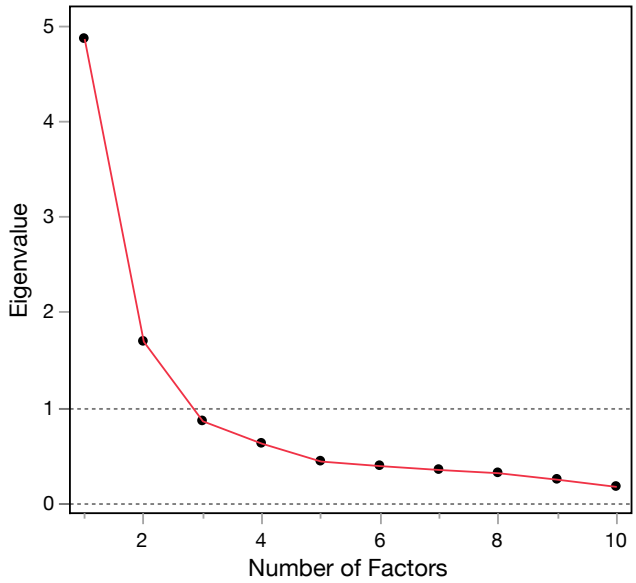
Factor	Variance	Percent	Cum Percent
Factor 1	2.8932	28.932	28.932
Factor 2	2.0153	20.153	49.085
Factor 3	1.4726	14.726	63.811

Significance Test

Test	DF	ChiSquare	Prob>ChiSq
H0: no common factors. HA: at least one common factor.	45	1229.759	<.0001*

Test	DF	Criterion	ChiSquare	Prob>ChiSq
H0: 3 factors are sufficient. HA: more factors are needed.	18	0.147	32.462	0.0194*

Scree Plot



Factor Analysis

Factor Analysis on Correlations with 3 Factors: Maximum Likelihood / Varim

Measures of Fit

Measures of Fit	Fit Index
Chi-Square without Bartlett's Correction	33.368
AIC	-2.632
BIC	-64.360
Tucker and Lewis's Index	0.969
Root Mean Square Error of Approximation	0.061

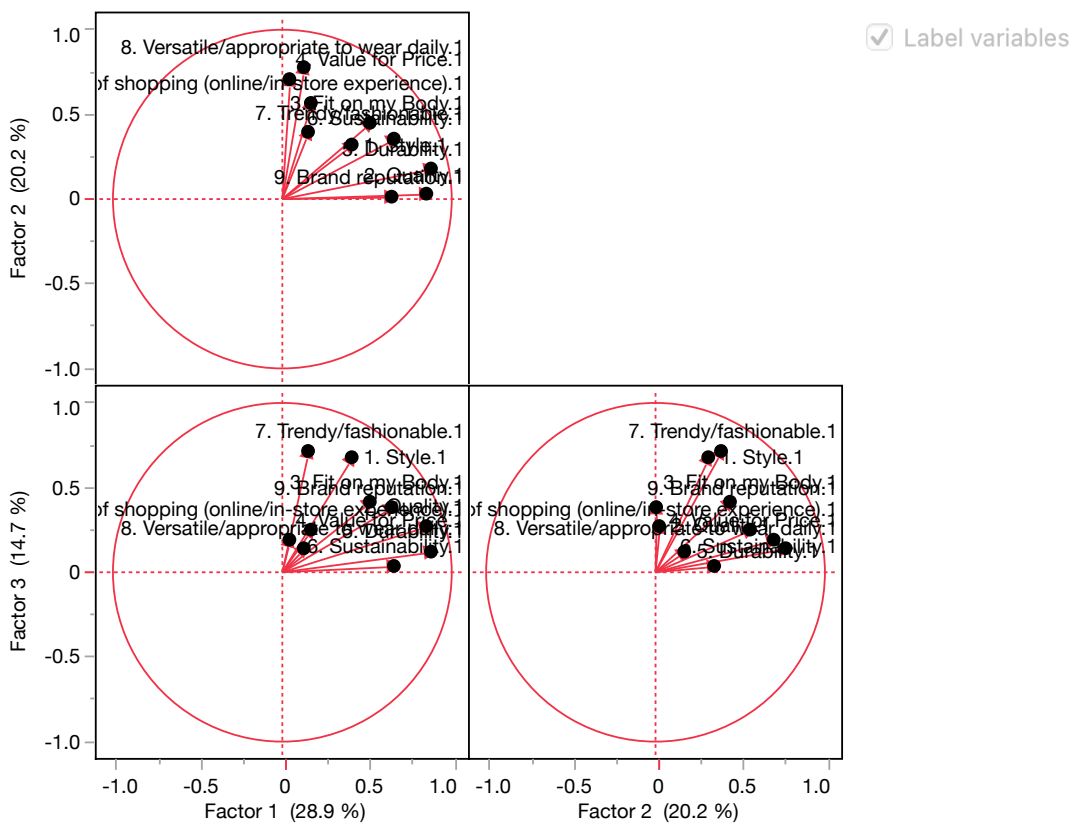
Rotated Factor Loading

	Factor 1	Factor 2	Factor 3
5. Durability.1	0.8827299	0.1750728	0.1164410
2. Quality.1	0.8553922	0.0267354	0.2673600
6. Sustainability.1	0.6637851	0.3513922	0.0319810
9. Brand reputation.1	0.6502781	0.0089583	0.3791964
3. Fit on my Body.1	0.5217900	0.4449069	0.4112107
8. Versatile/appropriate to wear daily.1	0.1317222	0.7735291	0.1366007
4. Value for Price.1	0.0470387	0.7028966	0.1869418
10. Ease of shopping (online/in-store experience).1	0.1738158	0.5622992	0.2464914
7. Trendy/fashionable.1	0.1570747	0.3919526	0.7107431
1. Style.1	0.4148519	0.3166378	0.6739001

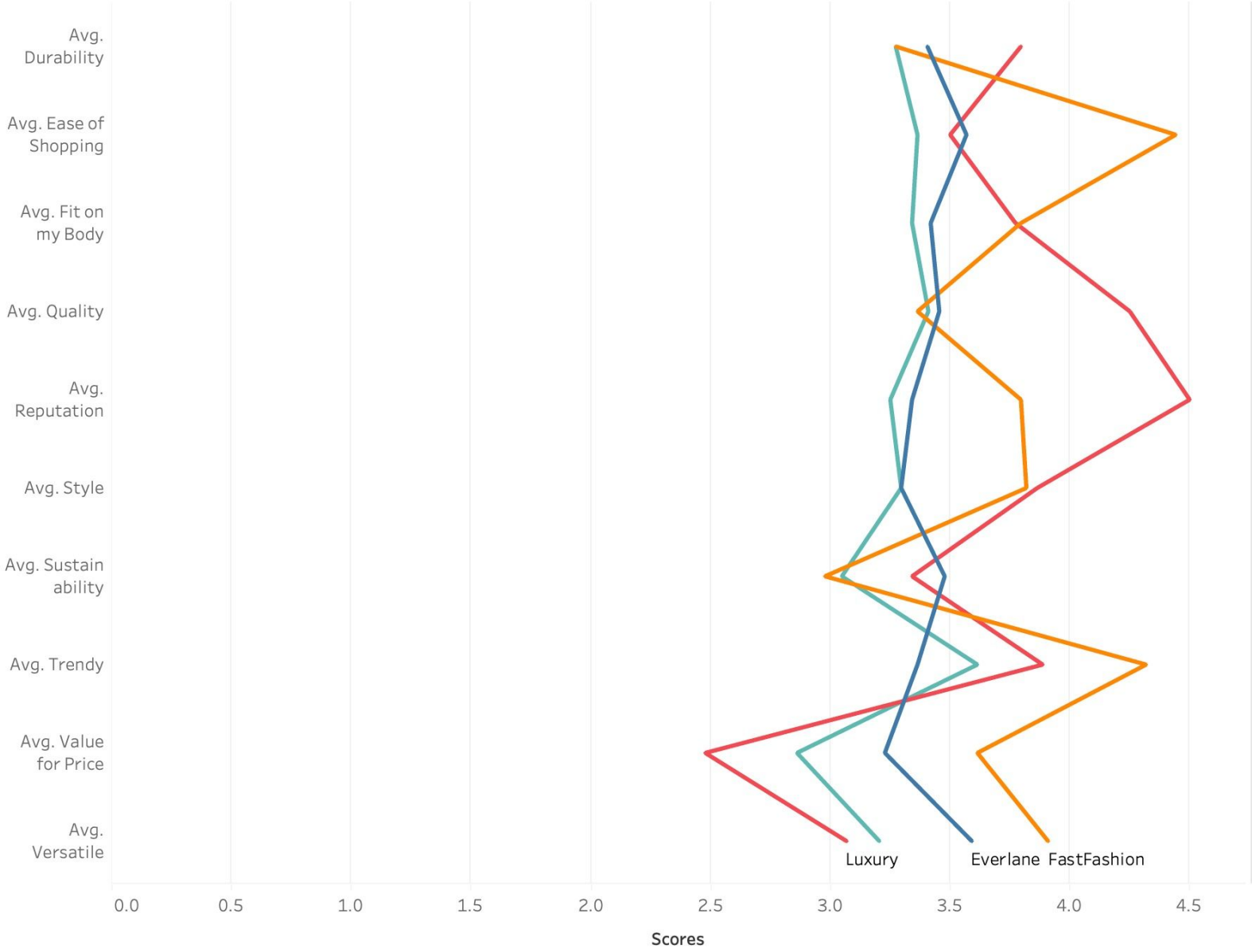
Suppress Absolute Loading Value Less Than 0.3

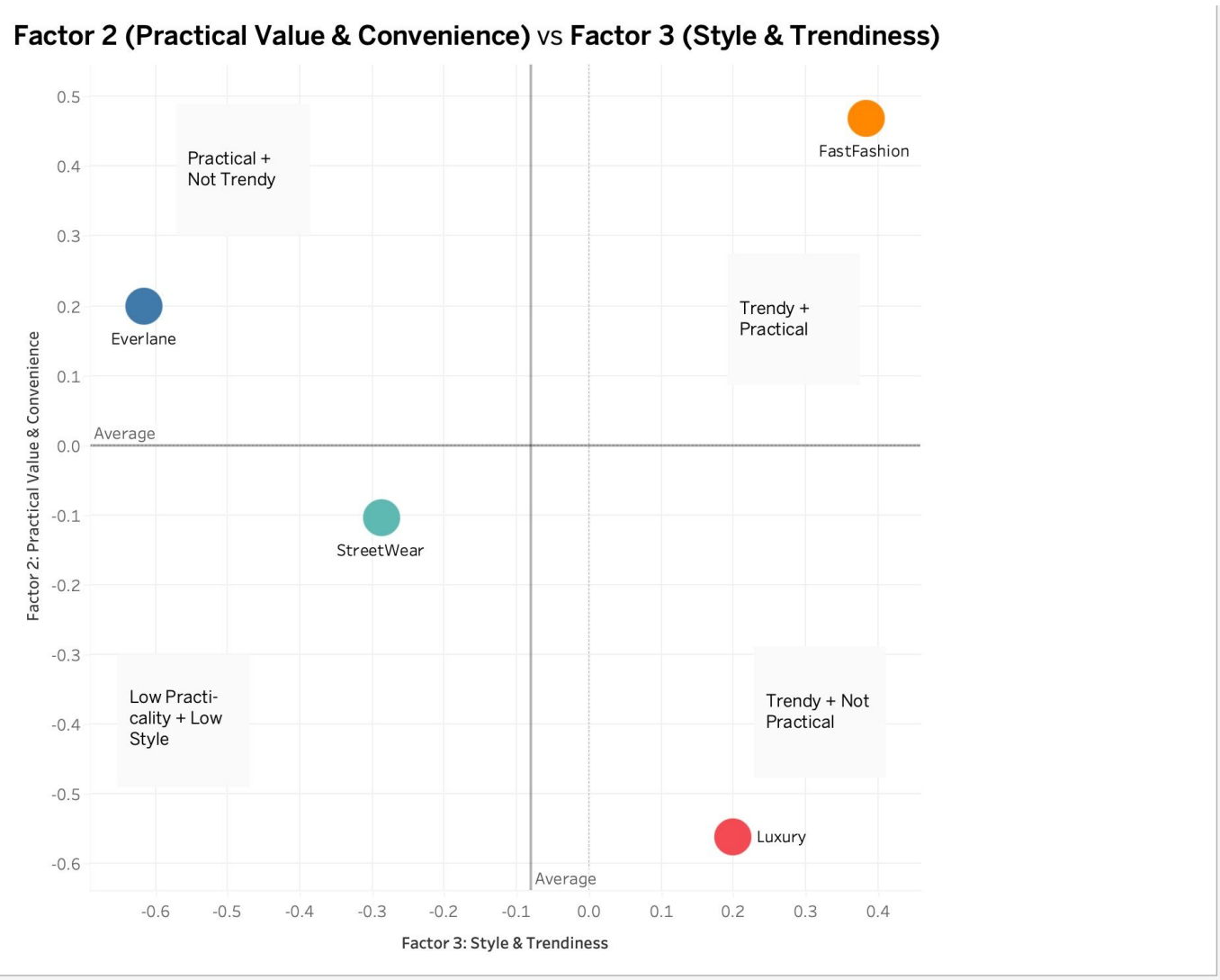
Dim Text 0.4

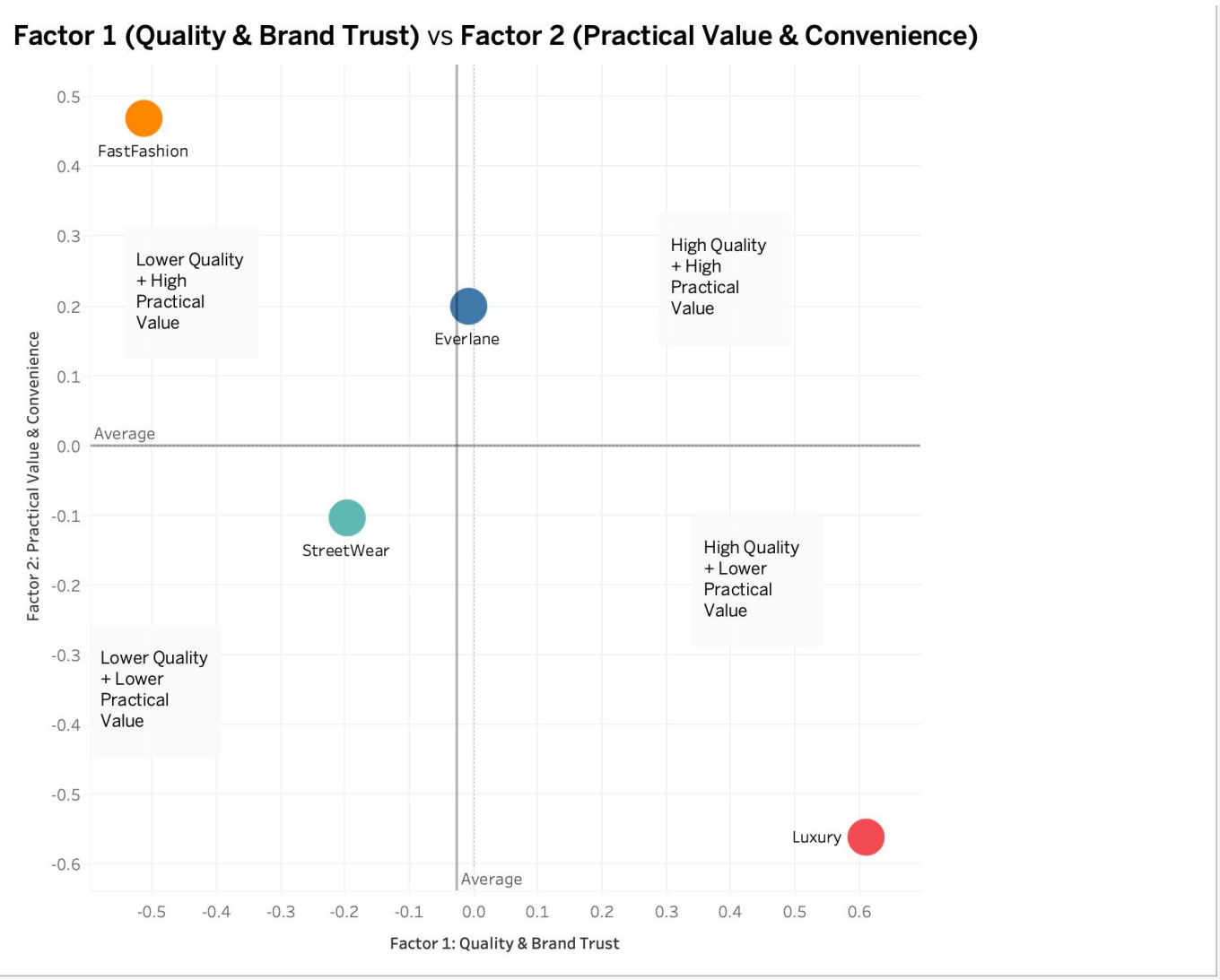
Factor Loading Plot

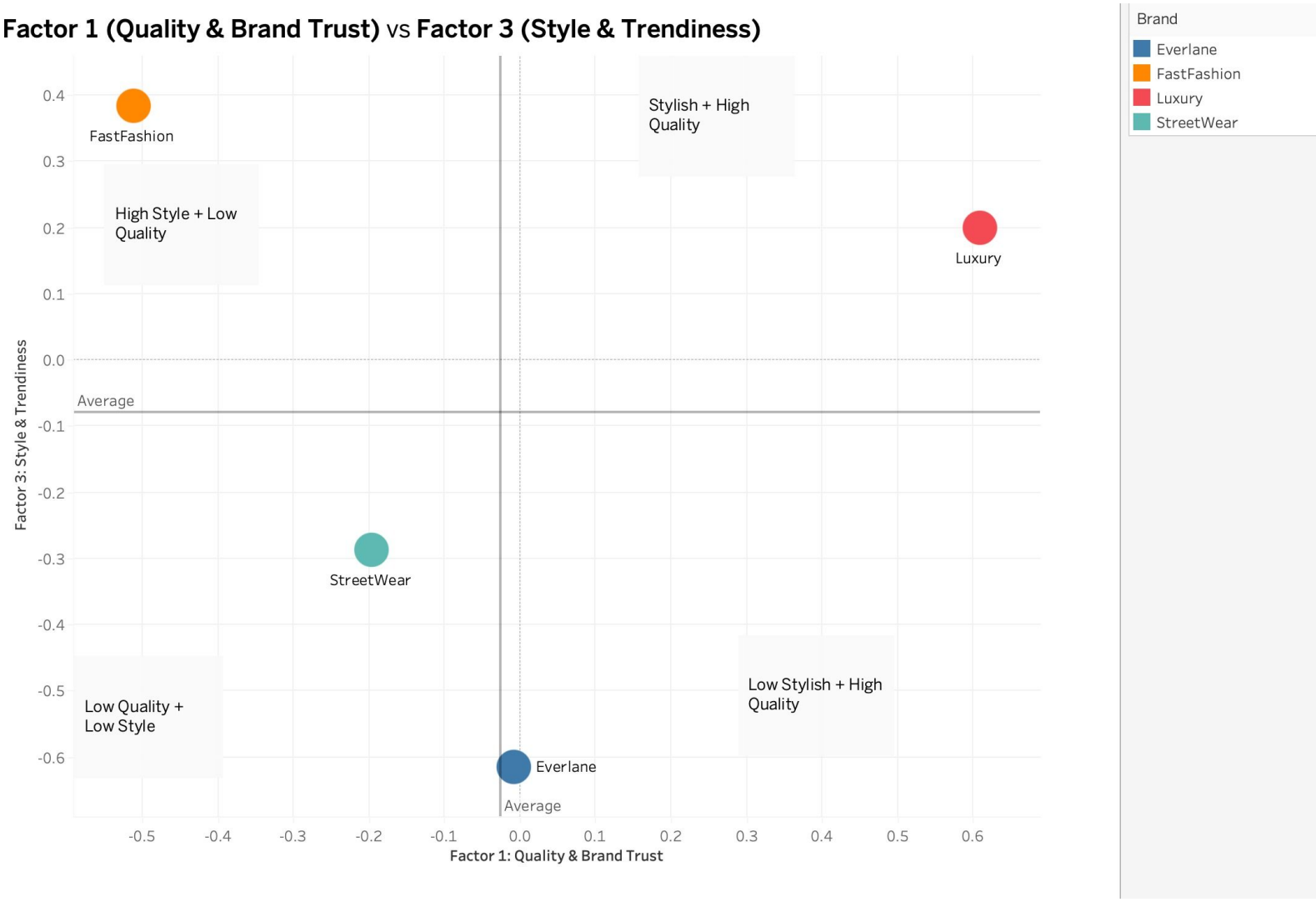


Snake Plot: Attribute-Level Brand Profiles Across 10 Dimensions




Factor 2 (Practical Value & Convenience) vs Factor 3 (Style & Trendiness)
Brand
Everlane
FastFashion
Luxury
StreetWear
Practical + Not Trendy
Trendy + Practical
Low Practi-cality + Low Style
Trendy + Not Practical
Average
Average
Factor 2: Practical Value & Convenience
Factor 3: Style & Trendiness
0.5
0.4
0.3
0.2
0.1
0.0
-0.1
-0.2
-0.3
-0.4
-0.5
-0.6
-0.6
-0.5
-0.4
-0.3
-0.2
-0.1
0.0
0.1
0.2
0.3
0.4


Factor 1 (Quality & Brand Trust) vs Factor 2 (Practical Value & Convenience)
Brand
Everlane
FastFashion
Luxury
StreetWear
FastFashion
Lower Quality + High Practical Value
High Quality + High Practical Value
Everlane
Average
StreetWear
High Quality + Lower Practical Value
Lower Quality + Lower Practical Value
Luxury
Average
Factor 2: Practical Value & Convenience
Factor 1: Quality & Brand Trust
0.5
0.4
0.3
0.2
0.1
0.0
-0.1
-0.2
-0.3
-0.4
-0.5
-0.6
-0.5
-0.4
-0.3
-0.2
-0.1
0.0
0.1
0.2
0.3
0.4
0.5
0.6


Factor 1 (Quality & Brand Trust) vs Factor 3 (Style & Trendiness)
Brand
Everlane
FastFashion
Luxury
StreetWear
Stylish + High Quality
High Style + Low Quality
Low Quality + Low Style
Low Stylish + High Quality
FastFashion
Luxury
StreetWear
Everlane
Average
Average
Factor 3: Style & Trendiness
Factor 1: Quality & Brand Trust
0.4
0.3
0.2
0.1
0.0
-0.1
-0.2
-0.3
-0.4
-0.5
-0.6
-0.5
-0.4
-0.3
-0.2
-0.1
0.0
0.1
0.2
0.3
0.4
0.5
0.6