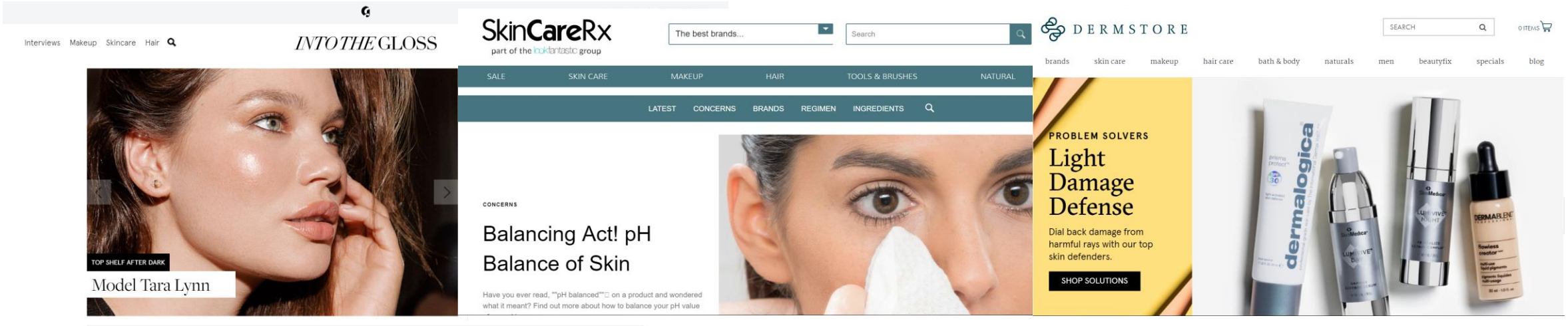
Correlation Analysis on Skincare Keywords

- Product Team
- Yuan Ding

Data Sources

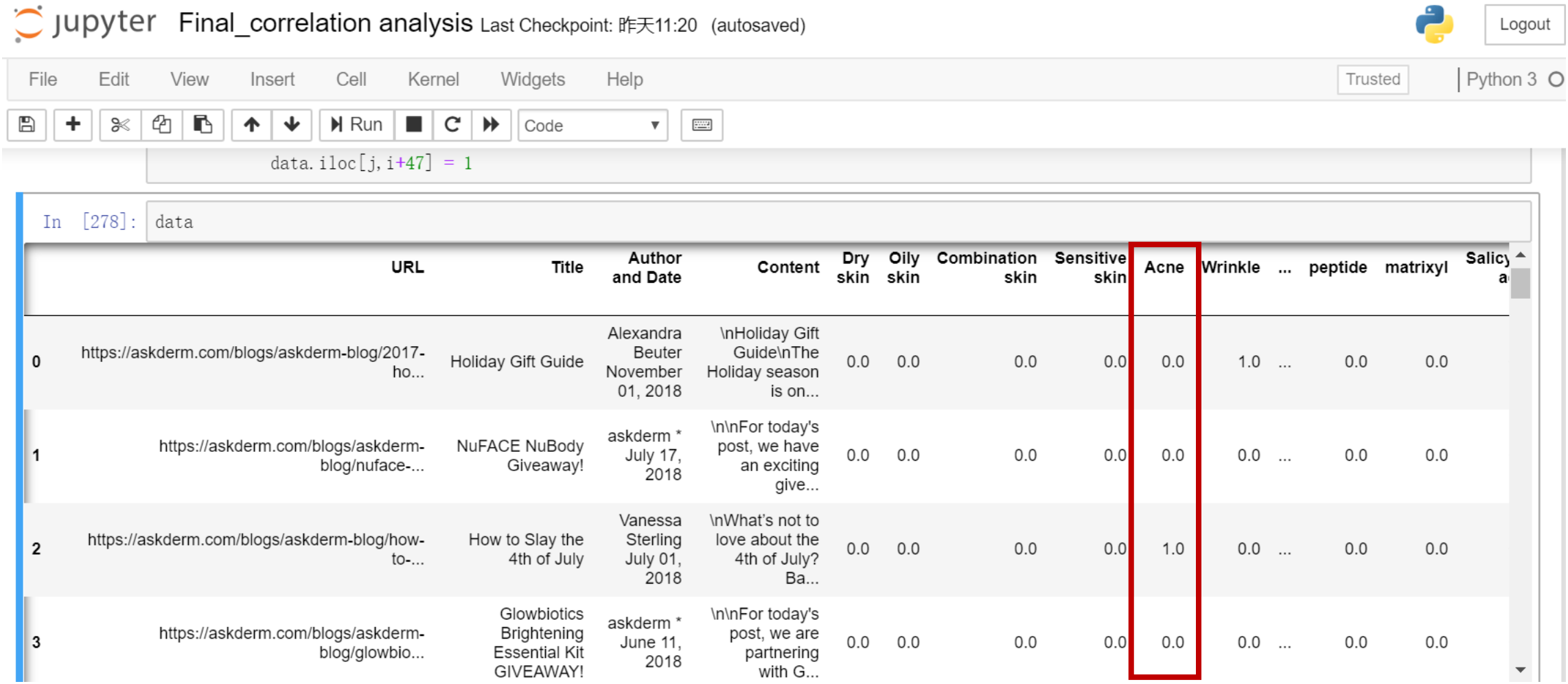
- **1811** articles from **5** main skincare blogs: Into The Gloss, Dermstore, b-glowing, Skinstore, Askderm.
- Year Range: recent **Five** years (2015-2019)
- Topic: all **skincare**



Method

- Use **web crawler** to retrieve the information (four columns: url, title, author and date, content) of each article in the website to form our **raw data frame**
- Analyze the text by checking if it contains a specific keyword in our keyword list (we have four kinds of keywords: **problems**, **products**, **season** and **ingredients**)
- Then conduct correlation analysis based on the **frequency** of occurrence of each keyword
- **Main packages:** Python (requests, Beautifulsoup, urllib, numpy, pandas)

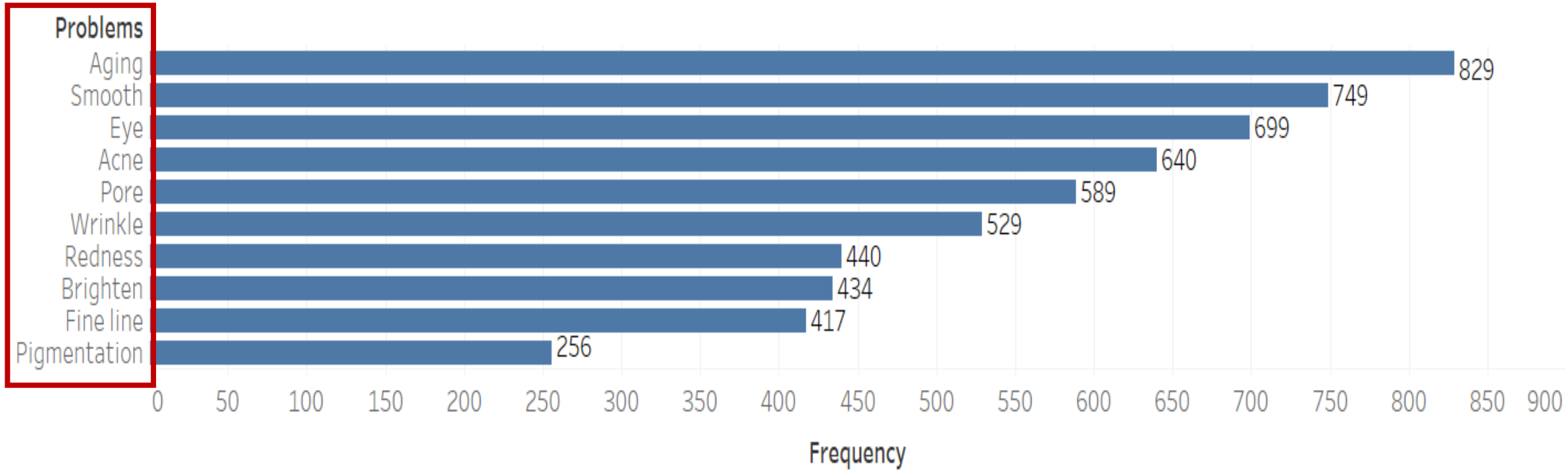
Final Data Frame



	URL	Title	Author and Date	Content	Dry skin	Oily skin	Combination skin	Sensitive skin	Acne	Wrinkle	...	peptide	matrixyl	Salicy a
0	https://askderm.com/blogs/askderm-blog/2017-ho...	Holiday Gift Guide	Alexandra Beuter November 01, 2018	\nHoliday Gift Guide\nThe Holiday season is on...	0.0	0.0	0.0	0.0	0.0	1.0	...	0.0	0.0	
1	https://askderm.com/blogs/askderm-blog/nuface-...	NuFACE NuBody Giveaway!	askderm * July 17, 2018	\n\nFor today's post, we have an exciting give...	0.0	0.0	0.0	0.0	0.0	0.0	...	0.0	0.0	
2	https://askderm.com/blogs/askderm-blog/how-to-...	How to Slay the 4th of July	Vanessa Sterling July 01, 2018	\nWhat's not to love about the 4th of July? Ba...	0.0	0.0	0.0	0.0	1.0	0.0	...	0.0	0.0	
3	https://askderm.com/blogs/askderm-blog/glowbio...	Glowbiotics Brightening Essential Kit GIVEAWAY!	askderm * June 11, 2018	\n\nFor today's post, we are partnering with G...	0.0	0.0	0.0	0.0	0.0	0.0	...	0.0	0.0	

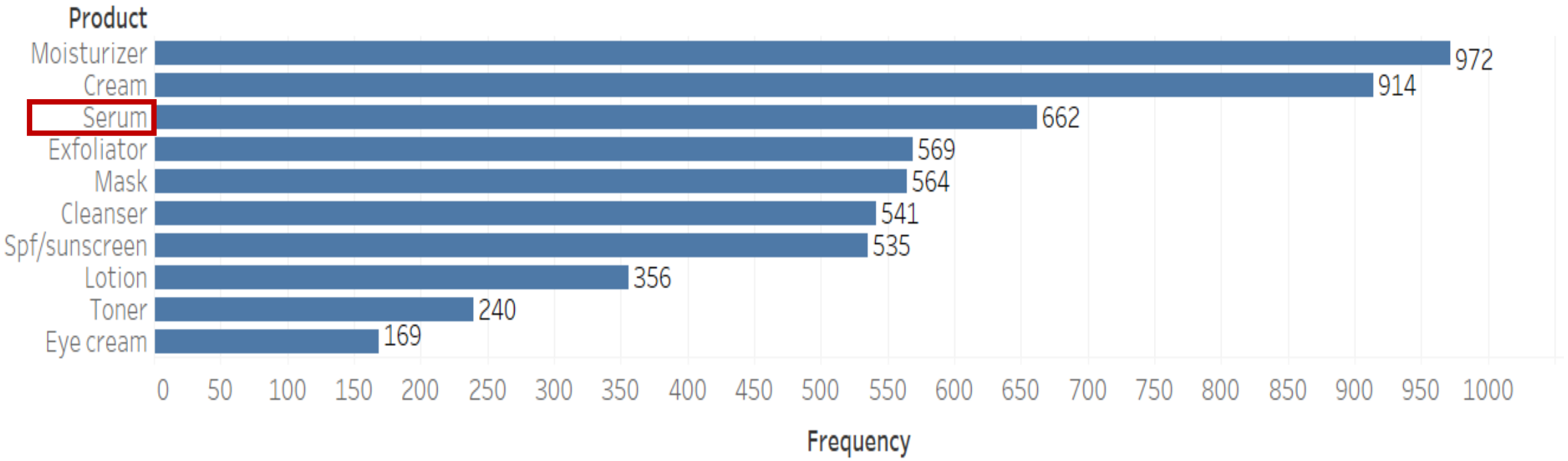
Part I: Univariate Analyses

- Top 10 skin problems/concerns



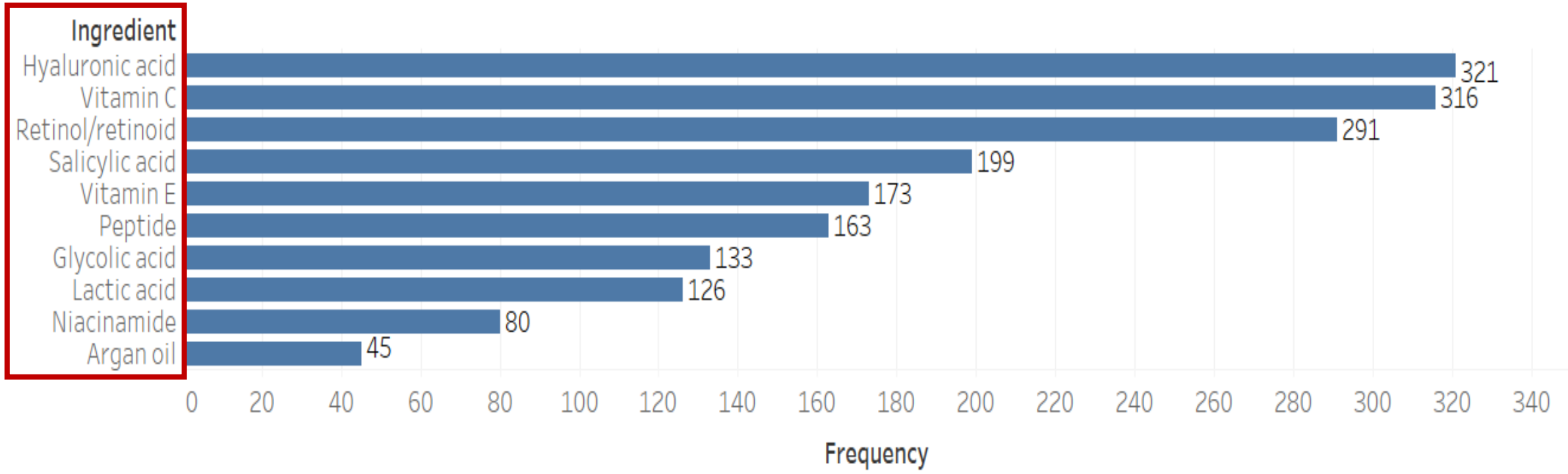
Part I: Univariate Analyses

- The most mentioned product types



Part I: Univariate Analyses

- Top 10 popular ingredients



Part I: Univariate Analyses

- Translation to products (1)
  - **Serum** is popular and highly mentioned these years (ranked #3, good to know!)
  - 5 functions our serum might have: **anti-aging**(wrinkles), **moisturizing**, **skin smoothing**(deal with acne/ clogged pores), **brightening**, **mild**(sensitive skin friendly, reduce redness)
  - Ingredients we may consider using for each function:
    Hyaluronic acid (moisturizing)
    Vitamin C (brightening, anti-aging)
    Retinol/retinoid (anti-aging)
    Salicylic acid (anti-acne, smoothing)

Part II: Bivariate Analyses (correlations)

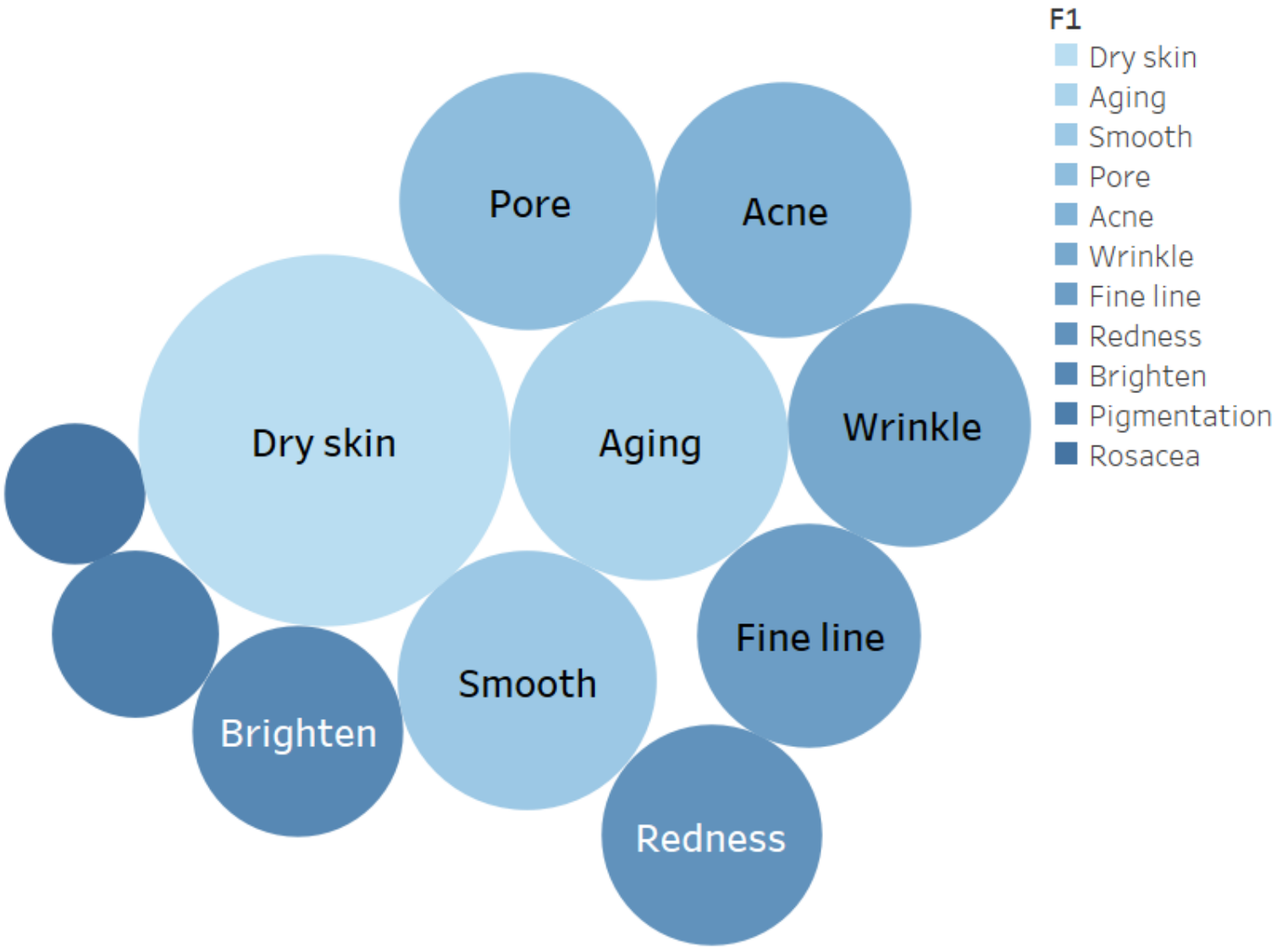
- Skin type & problems

	Smooth	Aging	Wrinkle	Fine line	Acne	Dark spot	Pigmentation	Blemish	Pore	Redness	Rosacea	Brighten	Firmness
Dry skin	0.486	0.565	0.428	0.363	0.473	0.128	0.202	0.130	0.479	0.353	0.144	0.322	0.051
Oily skin	0.557	0.597	0.453	0.343	0.667	0.149	0.244	0.254	0.667	0.313	0.149	0.348	0.040
Combination skin	0.638	0.550	0.450	0.388	0.588	0.138	0.188	0.288	0.713	0.450	0.163	0.338	0.088
Sensitive skin	0.500	0.583	0.409	0.338	0.520	0.129	0.242	0.172	0.465	0.381	0.187	0.338	0.073

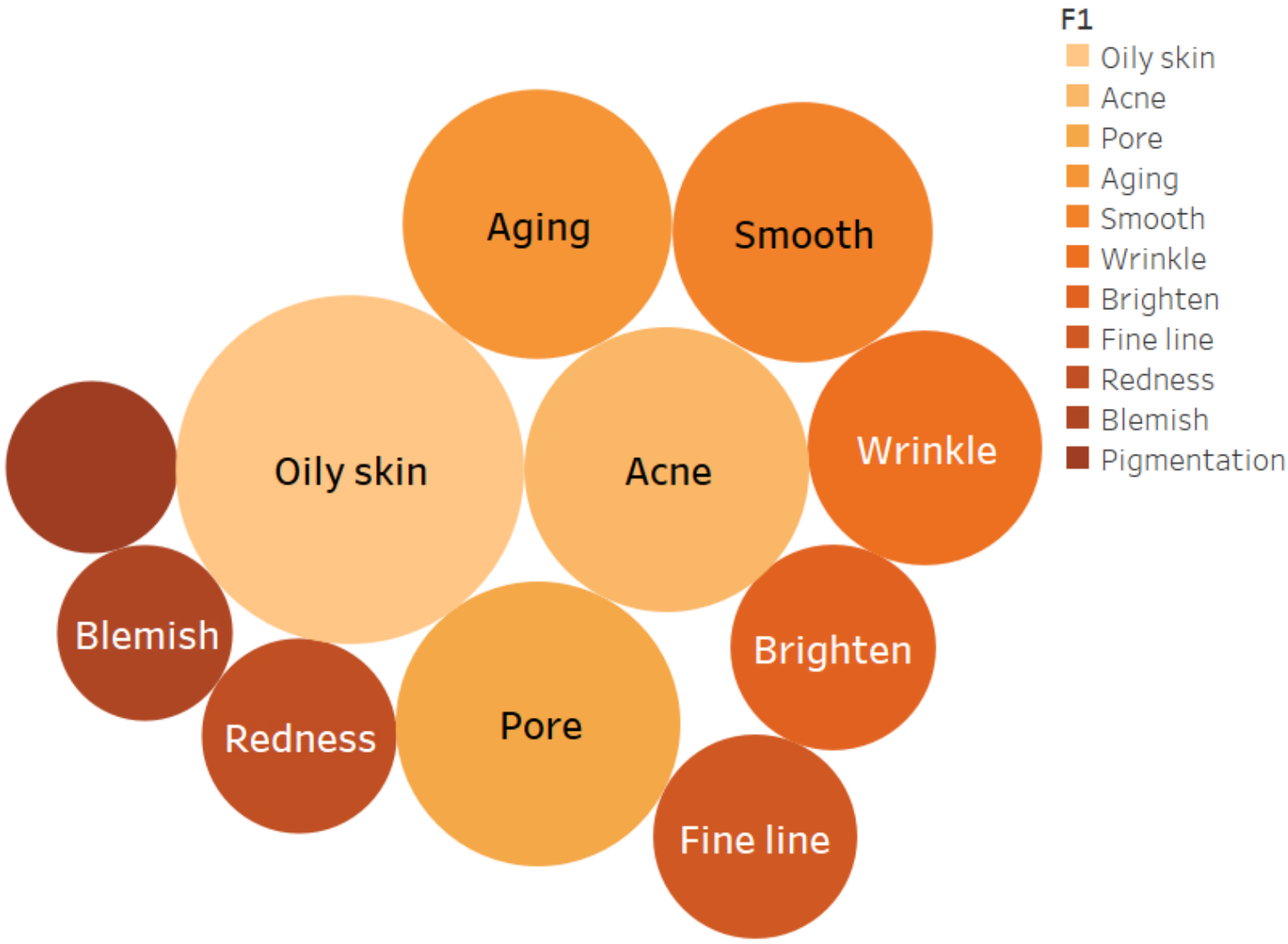
Part II: Bivariate Analyses (correlations)

- Define “correlation”
  - The “**correlation**” here refers to the percentage of articles containing another keyword in the articles containing a specific keyword. e.g. we have 292 articles containing the keyword “dry skin”, the correlation value “0.486” means that among all 292 articles, 48.6% of them also contain the keyword “smooth”.
  - It is not the exact definition in statistics!
  - In our analysis, we find out that if the “correlation” is higher than 0.4, then the two keywords can be seen as “**highly correlated**”.

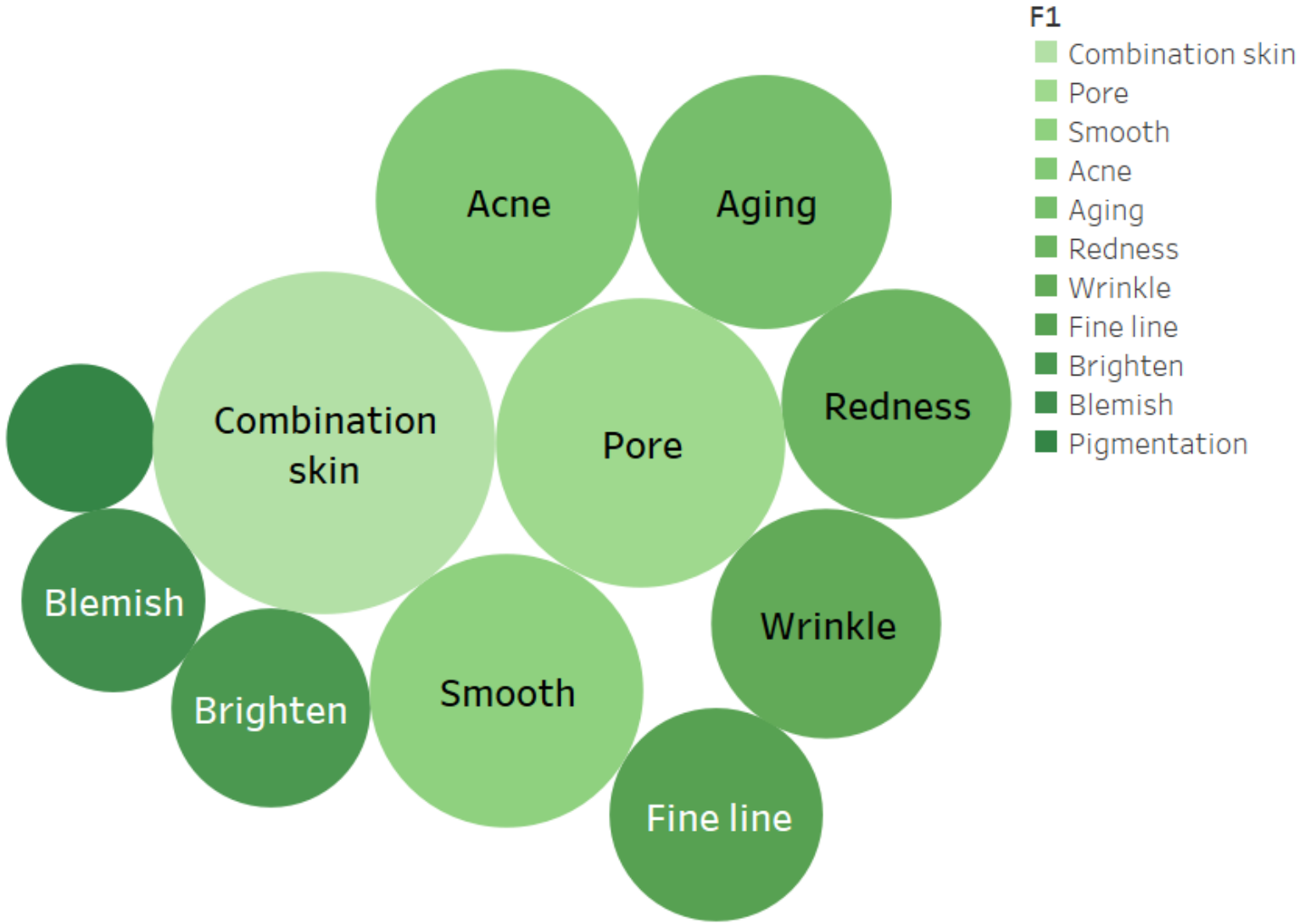
Top 10 Dry Skin Concerns



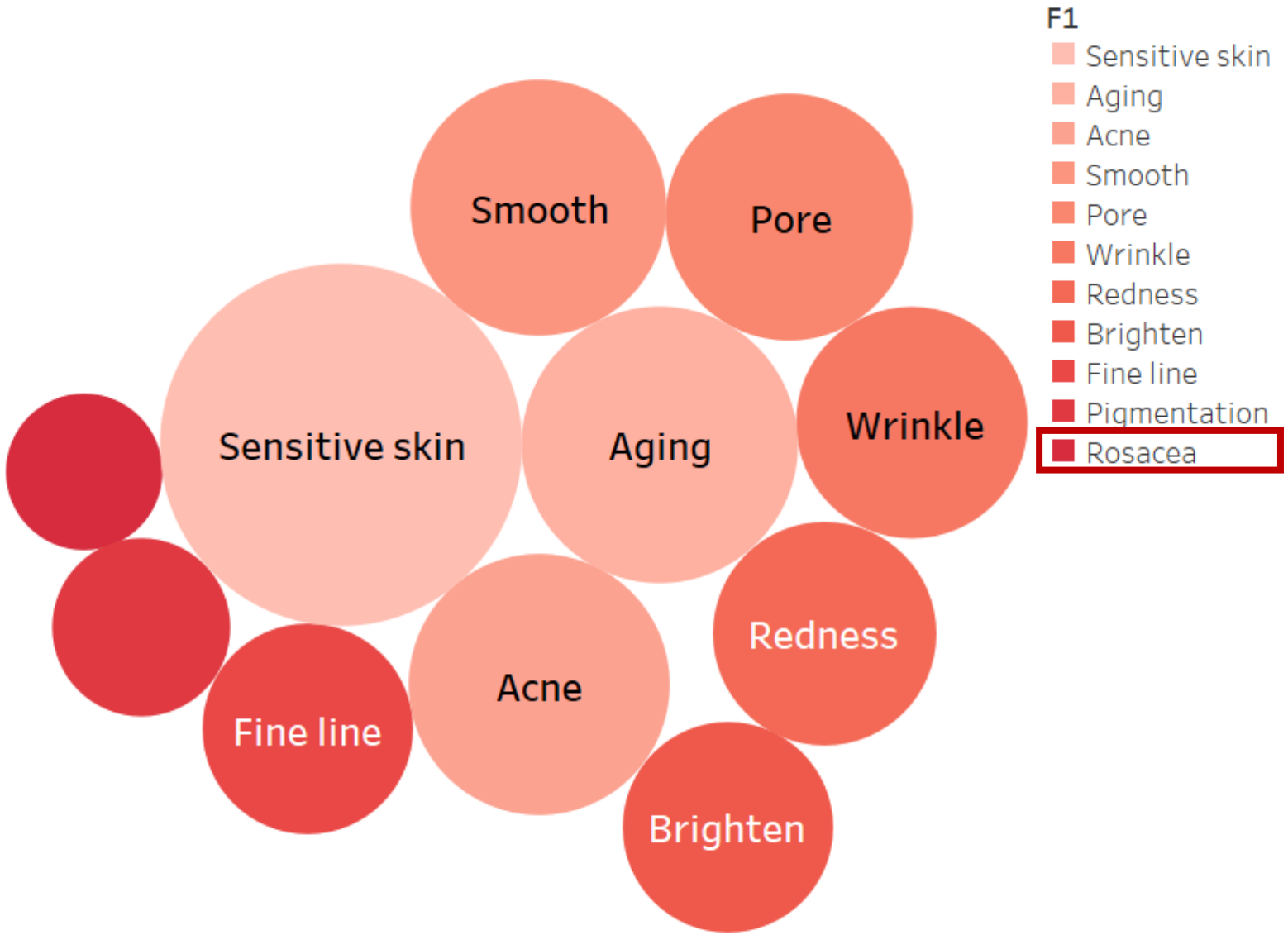
Top 10 Oily Skin Concerns



Top 10 Combination Skin Concerns



Top 10 Sensitive Skin Concerns



Translations (2)

- Some concerns are for all skin type : **aging**, **smoothness**, **acne**, **pore**, **wrinkle** (our product/skin report should emphasize these!)
- According to different skin type:
- **Oily/combination skin** is easier to have acne, blemish or pore problems
- **Combination/sensitive skin** is more likely to encounter redness/rosacea
- **Combination skin** cares more about firmness and smoothness
- **Sensitive skin** cares more about firmness

Part II: Bivariate Analyses (correlations)

- Skin type & products

	Cleanser	Exfoliator	Serum	Spf	Cream	Eye Cream	Lotion	Toner	Mask
Dry skin	0.455	0.428	0.473	0.353	0.606	0.110	0.257	0.195	0.366
Oily skin	0.547	0.517	0.537	0.403	0.567	0.119	0.269	0.274	0.408
Combination skin	0.475	0.400	0.525	0.475	0.675	0.175	0.313	0.238	0.388
Sensitive skin	0.485	0.442	0.442	0.396	0.596	0.131	0.278	0.179	0.331

All skin types use serum!

Part II: Bivariate Analyses (correlations)

- Season & problems/products

	Smooth	Aging	Wrinkle	Fine line	Acne	Dark spot	Pigmentation	Blemish	Pore	Redness	Rosacea	Brighten	Firmness
Spring	0.443	0.523	0.282	0.242	0.356	0.047	0.141	0.121	0.336	0.322	0.114	0.255	0.060
Summer	0.411	0.435	0.202	0.162	0.348	0.071	0.138	0.119	0.344	0.209	0.059	0.233	0.051
Autumn	0.399	0.535	0.253	0.167	0.409	0.121	0.182	0.152	0.374	0.278	0.056	0.293	0.030
Winter	0.438	0.456	0.257	0.190	0.279	0.058	0.137	0.093	0.314	0.221	0.084	0.252	0.018
Holiday	0.480	0.429	0.255	0.214	0.245	0.061	0.122	0.082	0.255	0.235	0.020	0.316	0.041
Travel	0.494	0.488	0.259	0.205	0.259	0.072	0.090	0.114	0.301	0.223	0.030	0.289	0.030

Part II: Bivariate Analyses (correlations)

- Season & products

	Cleanser	Exfoliator	Serum	Spf	Cream	Eye Cream	Lotion	Toner	Mask
Spring	0.389	0.302	0.443	0.456	0.591	0.101	0.302	0.148	0.369
Summer	0.304	0.308	0.387	0.490	0.609	0.079	0.364	0.142	0.364
Autumn	0.318	0.364	0.379	0.293	0.535	0.131	0.177	0.162	0.409
Winter	0.358	0.345	0.425	0.381	0.646	0.102	0.296	0.150	0.385
Holiday	0.316	0.316	0.388	0.235	0.510	0.112	0.224	0.071	0.378
Travel	0.295	0.404	0.367	0.295	0.536	0.151	0.289	0.175	0.458

Translations (3)

- We didn’t see a significant difference in the product use or preference, although there is some slight difference.
- **Dry skin** uses **cream** while **oily skin** uses more **cleanser**, **exfoliator** and **toner**, which is much lighter for oily skin.
- Serum is used among **all skin types**, **all seasons**, only the emphasis will be slightly different.
- **Spring** is relatively higher correlated with **redness** and **rosacea** (pollen or seasonal change)
- We easily have **dark spots/pigmentation** problems in autumn, and we need **brightening** (bad sunlight in summer)
- In **spring/summer** the product emphasize **SPF** while in **autumn/winter** we emphasize **moisturizing**

Part II: Bivariate Analyses (correlations)

- Problems & ingredients

	Hyaluronic acid	Vitamin C	Retinol	Salicylic acid	Vitamin E	Peptide	Glycolic acid	Lactic acid	Niacinamide	Argan oil
Aging	0.273	0.264	0.222	0.124	0.141	0.157	0.090	0.098	0.068	0.034
Smooth	0.278	0.239	0.208	0.132	0.140	0.139	0.124	0.111	0.060	0.025
Acne	0.236	0.230	0.263	0.248	0.116	0.097	0.130	0.128	0.067	0.038
Wrinkle	0.346	0.314	0.293	0.146	0.157	0.221	0.129	0.108	0.076	0.030
Fine line	0.367	0.374	0.309	0.149	0.182	0.228	0.122	0.122	0.079	0.034
Dark spot	0.303	0.467	0.441	0.243	0.125	0.171	0.138	0.145	0.138	0.033
Pigmentation	0.336	0.430	0.367	0.242	0.184	0.160	0.184	0.176	0.117	0.039
Blemish	0.204	0.274	0.296	0.317	0.161	0.148	0.135	0.170	0.100	0.043
Pore	0.236	0.200	0.233	0.234	0.109	0.097	0.151	0.132	0.059	0.037
Redness	0.259	0.236	0.280	0.191	0.105	0.127	0.102	0.125	0.080	0.027
Rosacea	0.305	0.240	0.279	0.208	0.097	0.143	0.084	0.123	0.123	0.019
Brighten	0.295	0.403	0.290	0.194	0.159	0.184	0.143	0.150	0.094	0.035
Firmness	0.500	0.390	0.310	0.140	0.220	0.390	0.160	0.110	0.110	0.010

Translations (3)

- Function & Ingredients
  - **Acne**: Salicylic acid
  - **Brightening**: Vitamin C
  - **Wrinkle/fine line**: hyaluronic acid, vitamin C, retinol
  - **Dark spot/pigmentation**: hyaluronic acid, vitamin C, retinol
  - **Pore**: hyaluronic acid， retinol
  - **Mature**: peptide (matrixyl)
  - **Firmness**: hyaluronic acid, peptide

  - We should add **guide** to the ingredients (e.g. function/potential **risk**/limitation)

Next: Personalization(Go deeper!)

Sources

- Into the Gloss skincare: https://intothegloss.com
- SkinStore skincare: https://www.skinstore.com/blog/skincare/
- Dermstore: https://www.skinstore.com/blog/skincare/
- B-glowing: https://www.b-glowing.com/blog/category/skincare/
- Askderm: https://askderm-blogaskderm.com/blogs/