



Deeper **Discovery** Report

Dawei Jia | Cenli Han | Siqi Zhi | Jinchuan Yang | Yufei Jiang

Part 1

Segmentation and exploratory analysis
using PCA, clustering and Lasso

Segmentation



Using PCA (principle component analysis), we find PC1 and PC2 can explain more than 60% of variance. Then, we use PC1 and PC2 to do clustering to segment two different groups of customers. The characters of two segments are:

Segment 1: Big Customers

The median employee is 26418; the median monthly page volume is 4,673,326; the median machine number is 607

Segment 2: Small Customers

The median employee is 1670; the median monthly page volume is 230,965; the median machine number is 15

Recommend Score of Segments

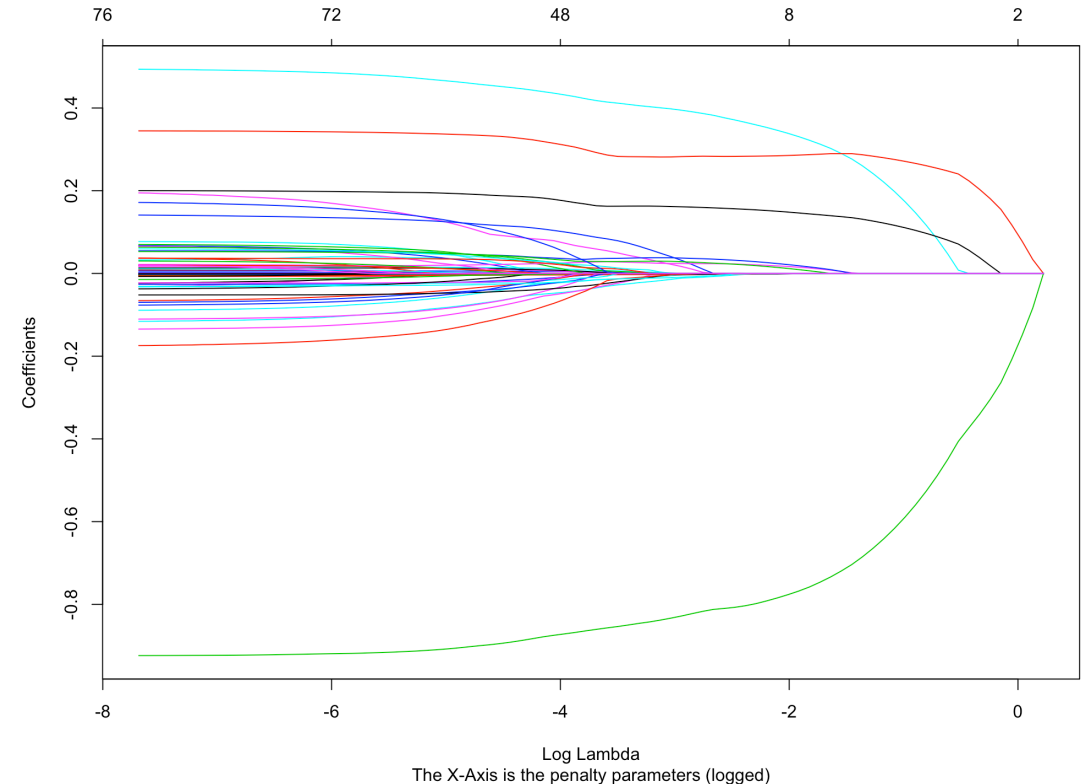
	Count	Count of Detractor	Count of Passive	Count of Promoter	Median of Recommend
Big customers	501	300	318	536	8
Small customers	388	294	84	103	5

The reason why we use median instead of average to explore data is outliers will influence the result of mean.

We found big customers has more promoters than detractors, while small customers has more detractors than promoters. Meanwhile, big customers has higher recommend score than small customers.

Big customers: Key Drivers of Recommend

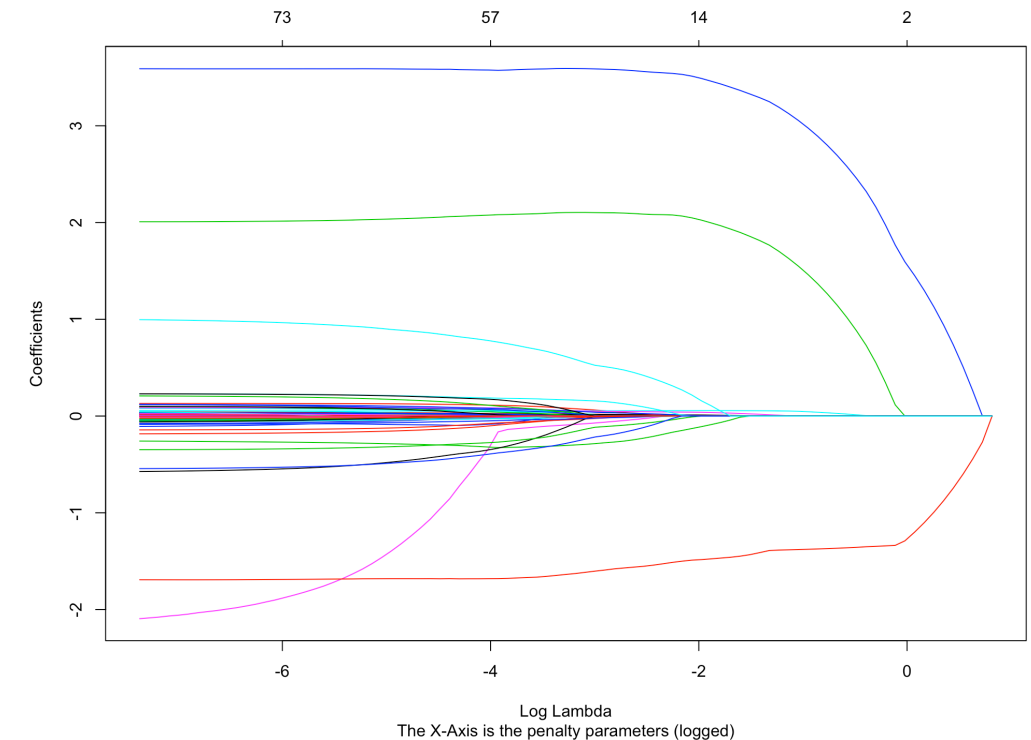
Promoter	0.36
Overall Equipment	0.28
Overall Sales	0.15
Overall Supplies	0.03
Repurchase	0.02
Drilldown: Average of Tech Service - Remote Support	0.02
Detractor	-0.80



According to lasso coefficients, we found that **overall equipment** and **overall sales** are two factors which have biggest impact on recommend score.

Small customers: Key Drivers of Recommend

Drilldown: Average of Supplies - Ease of orders	0.09
Overall Sales	0.05
Overall Invoicing	0.03
2016 - Telemarketing Call	0.02
Drilldown: Average of INVOICING - Timely	-0.07
Drilldown2: Average of IINVOICING - Ease of Help`	-0.12
016 - Action Required : Dissatisfied / Bounced Email	-0.23

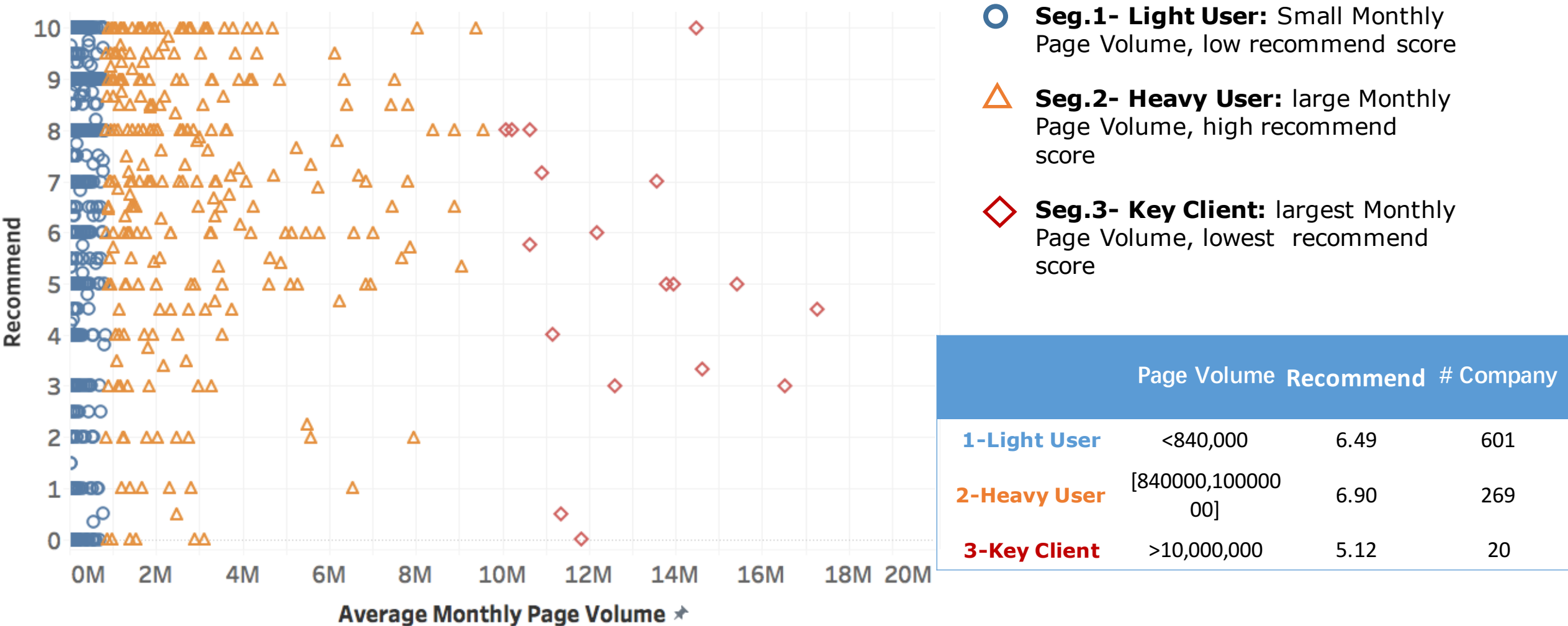


According to lasso coefficients, we found that positive influencers' coefficients are small. But **dissatisfied emails** and **ease of help** are two factors which prevent recommend score.

Part 2

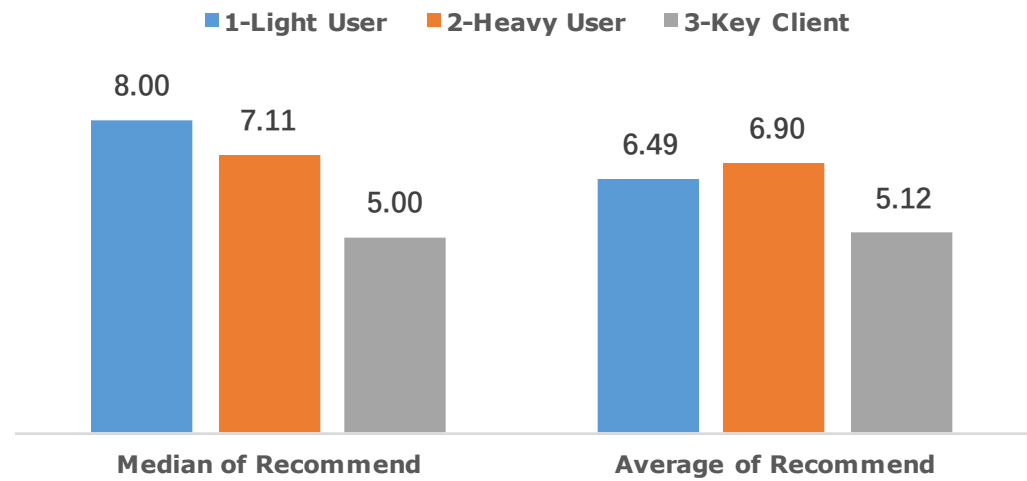
Cluster Segmentation and exploratory analysis only based on page volume and recommend score

Three segments based on **Page Volume** and **Recommend** via cluster



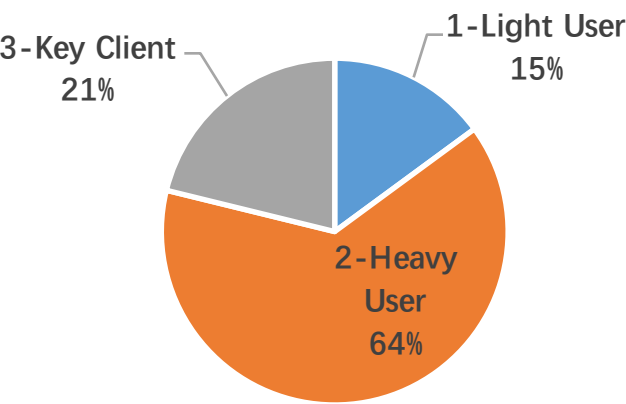
Closer look at clients in three segments

Recommend score distribution

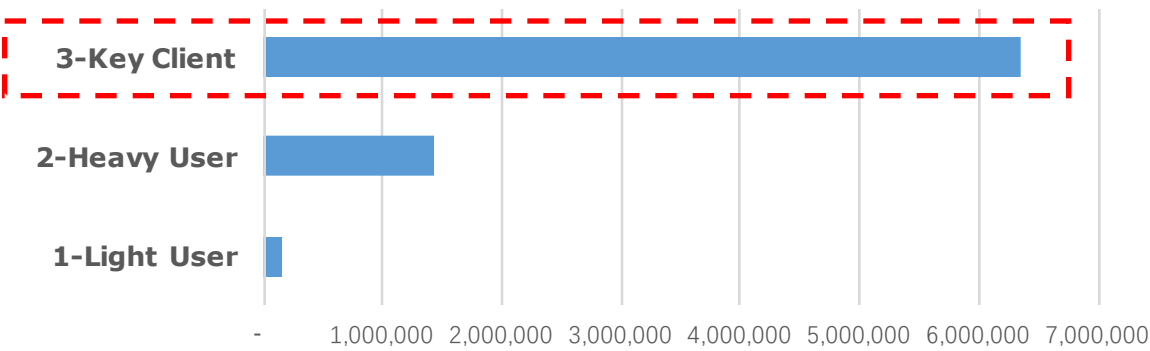


	Median of Recommend	Average of Recommend
1-Light User	8.00	6.49
2-Heavy User	7.11	6.90
3-Key Client	5.00	5.12

Revenue Constitution

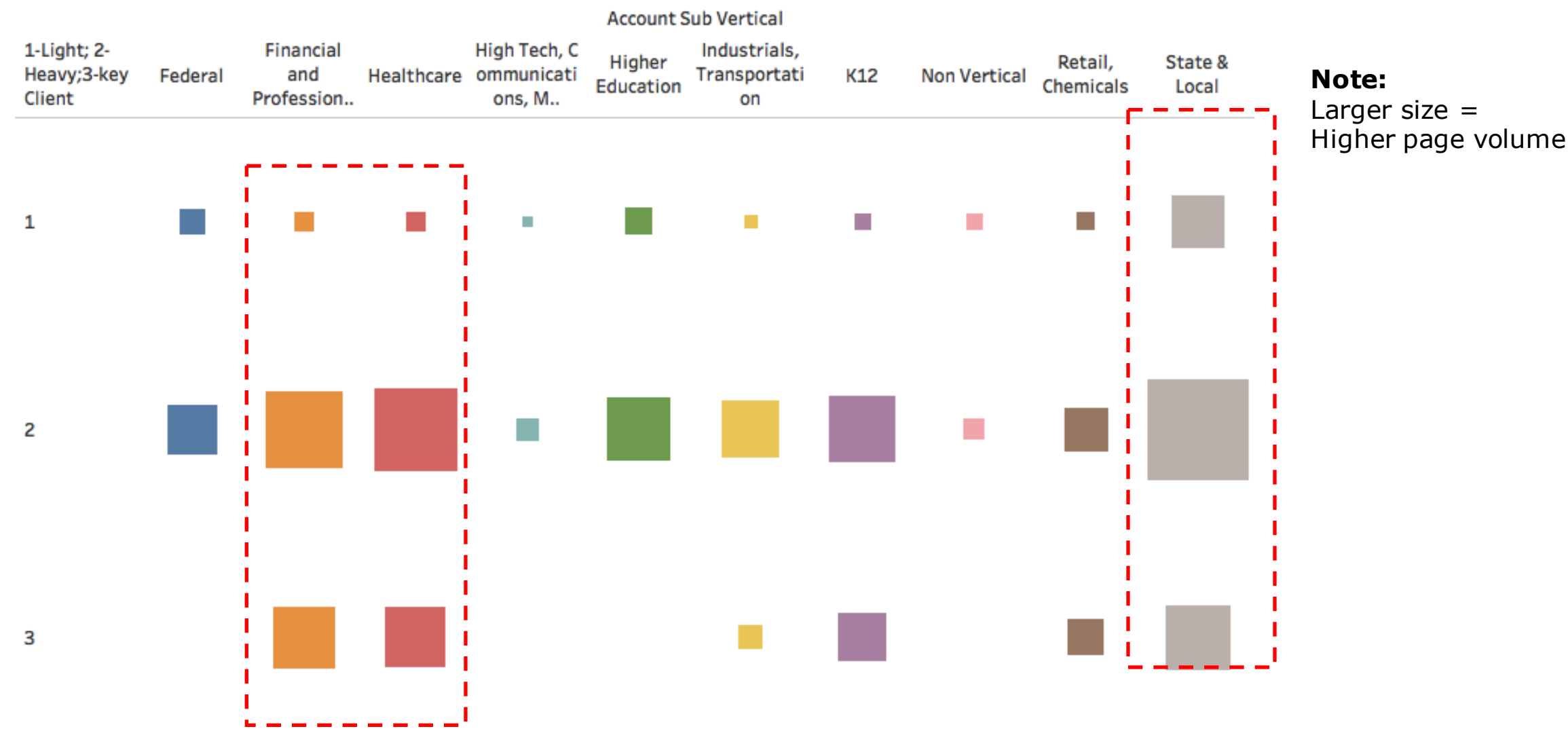


Average Revenue per Client

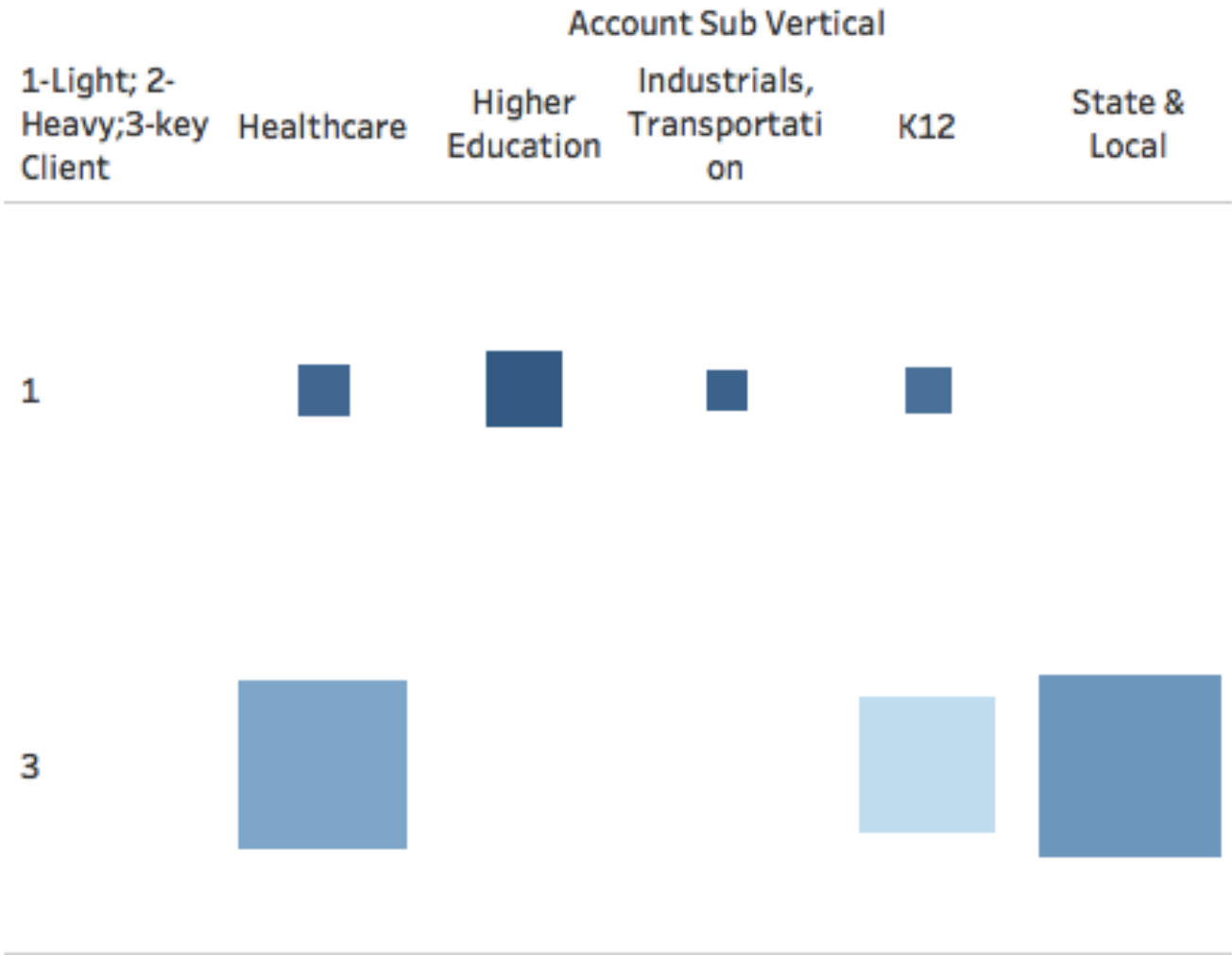


	1-Light User	2-Heavy User	3-Key Client
Average Revenue per Client	148,676	1,423,129	6,340,339

Closer look at clients by industry in three segments



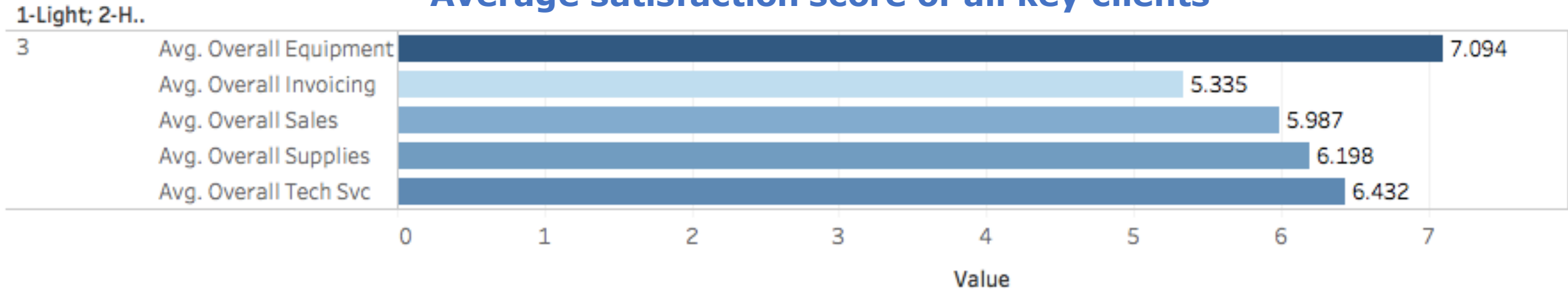
Closer look at “detractor” clients by industry in three segments



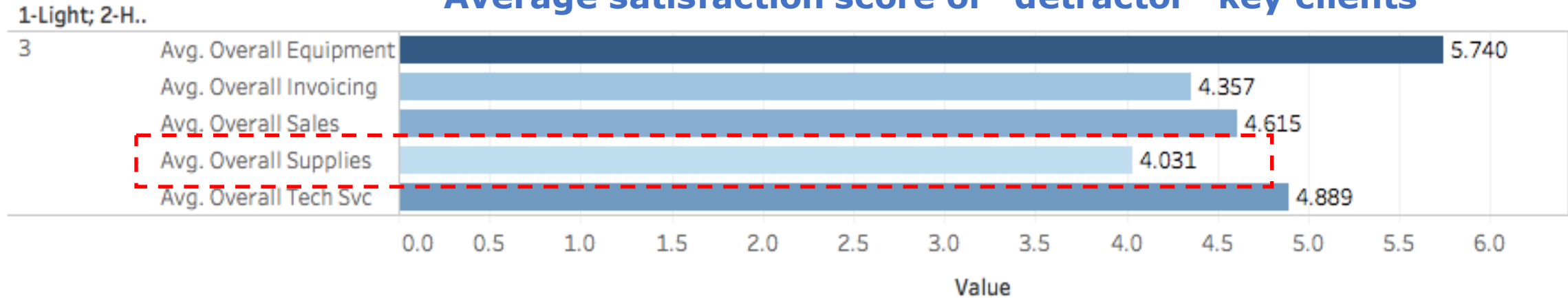
Note:
Larger size = Higher page volume;
lighter blue = lower recommend score

For **key clients** with low recommend score, the major problem is overall supplies

Average satisfaction score of all key clients

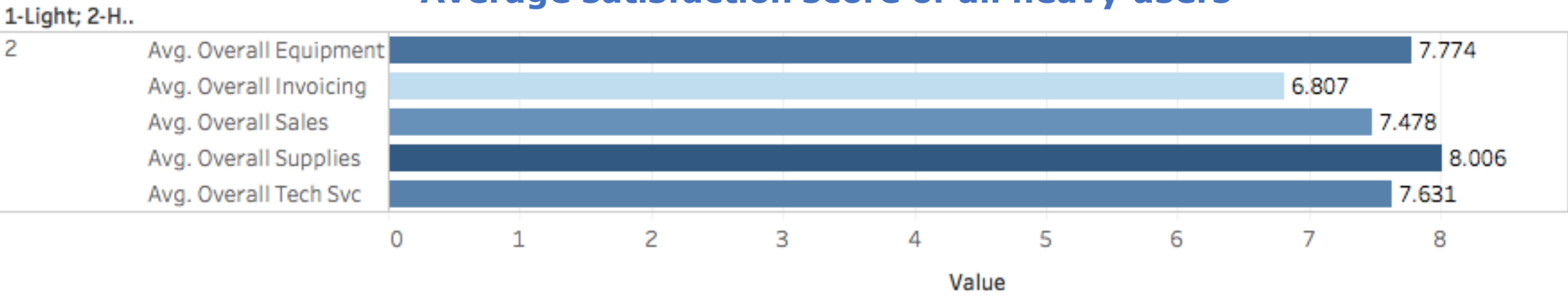


Average satisfaction score of "detractor" key clients

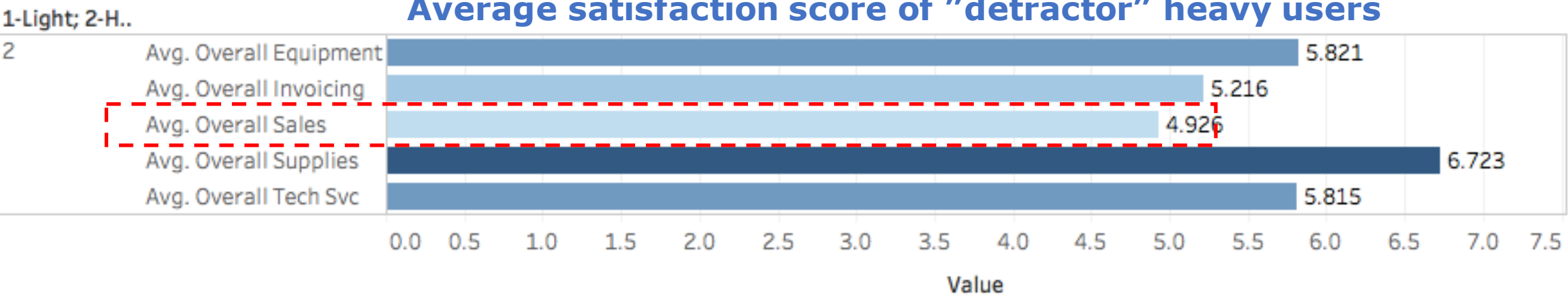


For heavy users, improvement potential is overall sales

Average satisfaction score of all heavy users

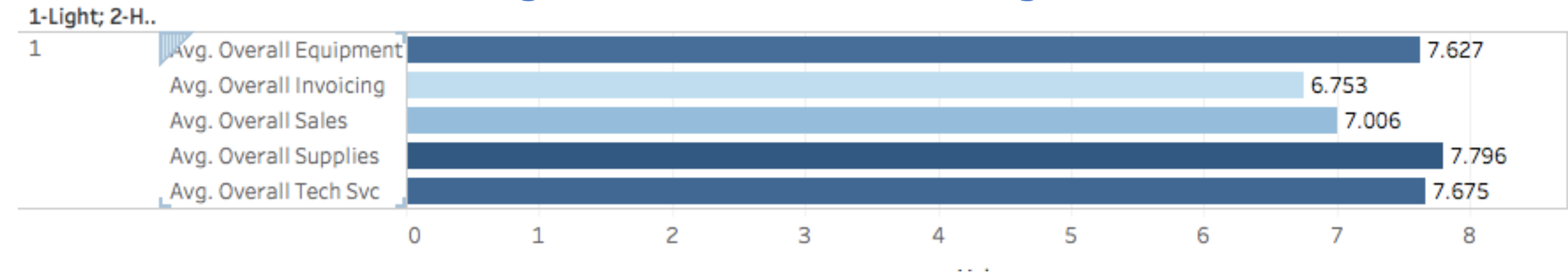


Average satisfaction score of "detractor" heavy users



For **light users**, major reason for detractors are overall invoicing and sales

Average satisfaction score of all light users



Average satisfaction score of "detractor" light users

