Market and customers data analysis

for a programming online learning company

Analyst: Canfang Feng

Currently, there are many programming online learning resources. These sources are intendent of each other.

On average, customers use at least three and most of them prefer the free resources.

online learning resources
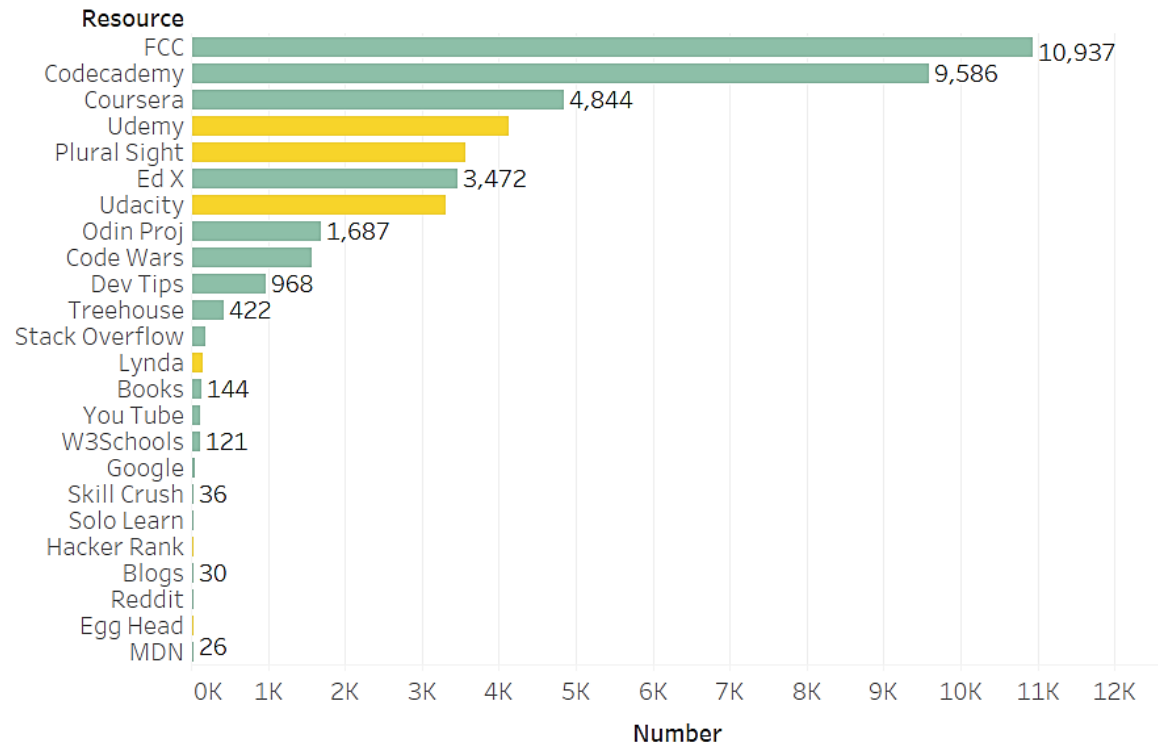
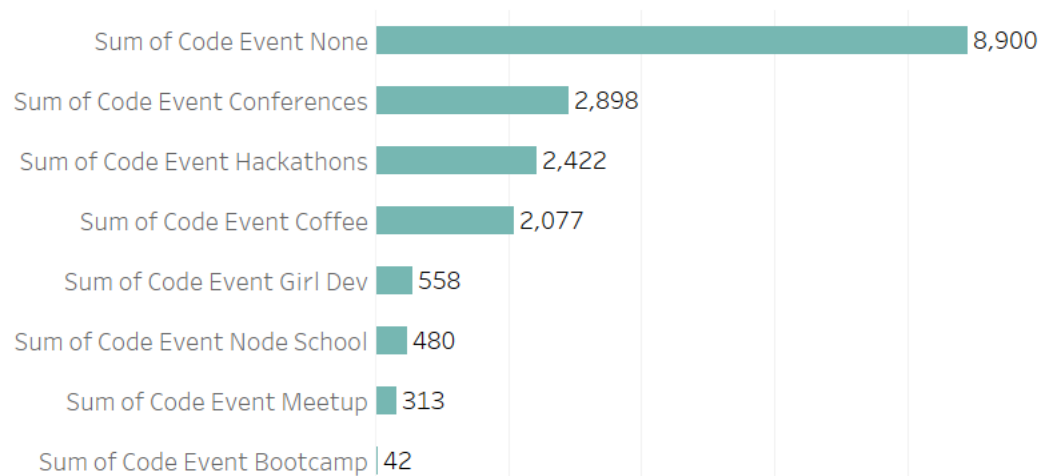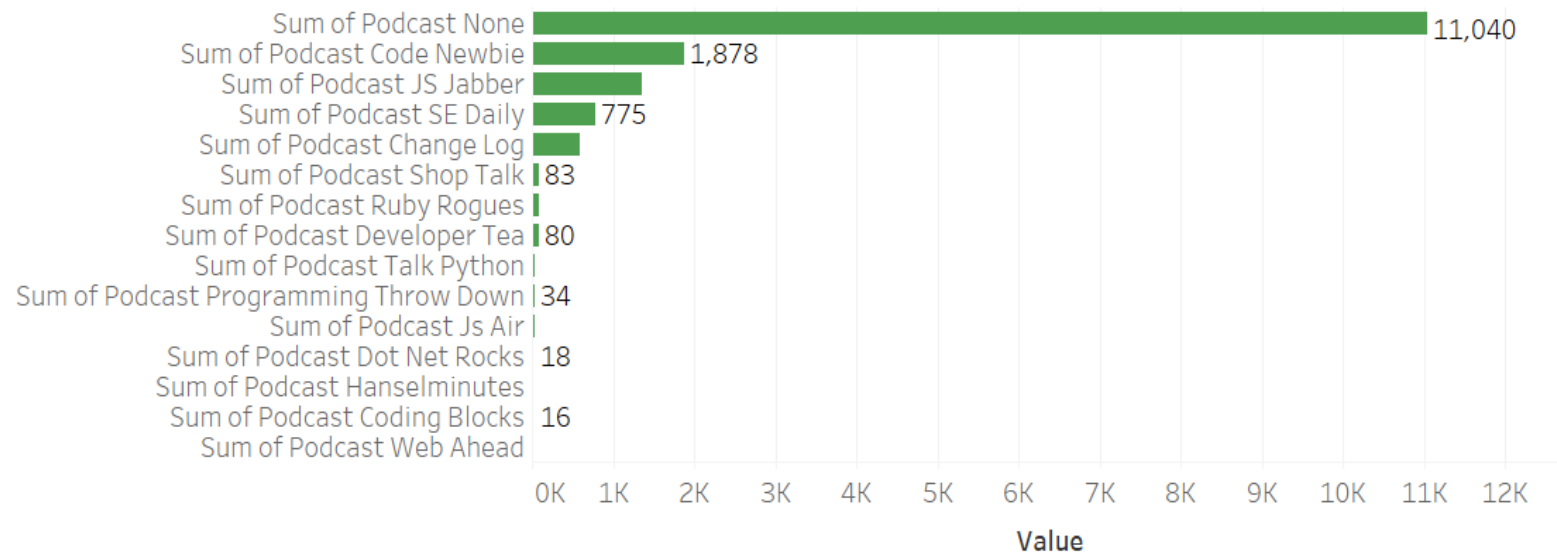Number as an attribute for each Resource. Color shows details about Free.

1 market overview

1.1 competitive landscape

Most of them didn't attend any in person coding events or listen coding-related podcasts.

EVENT

Event	Value
Sum of Code Event None	8,900
Sum of Code Event Conferences	2,898
Sum of Code Event Hackathons	2,422
Sum of Code Event Coffee	2,077
Sum of Code Event Girl Dev	558
Sum of Code Event Node School	480
Sum of Code Event Meetup	313
Sum of Code Event Bootcamp	42

coding-related podcasts

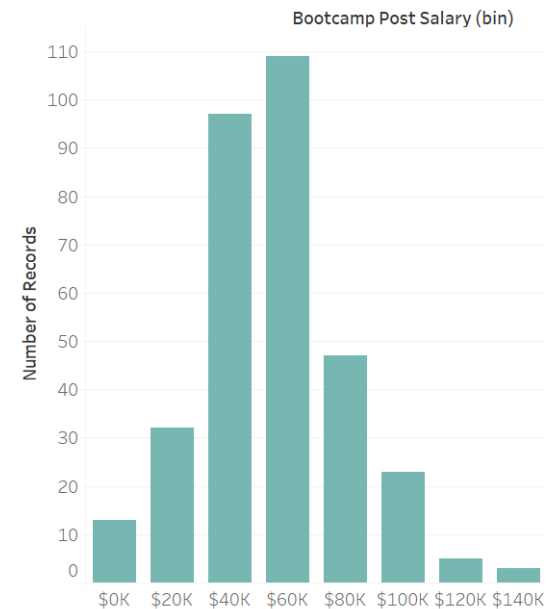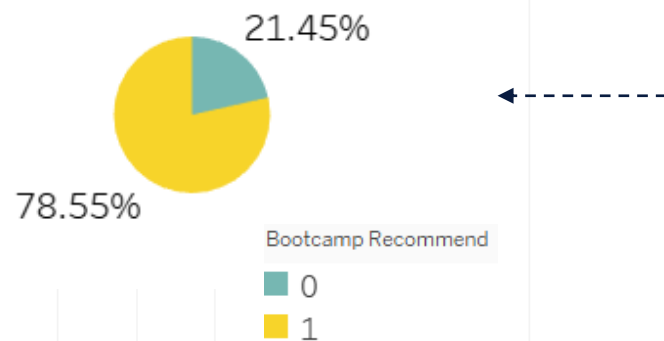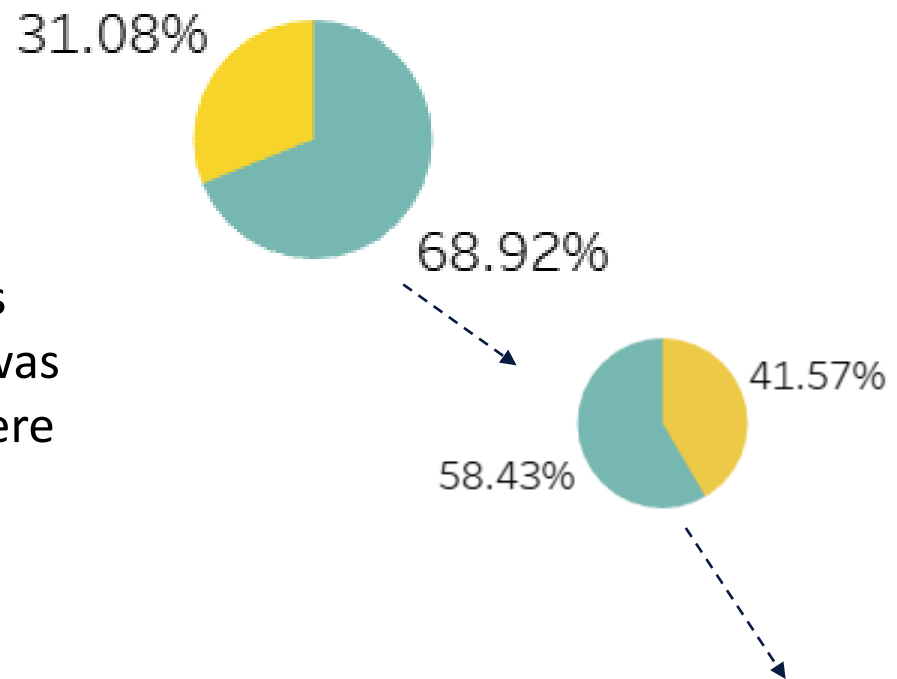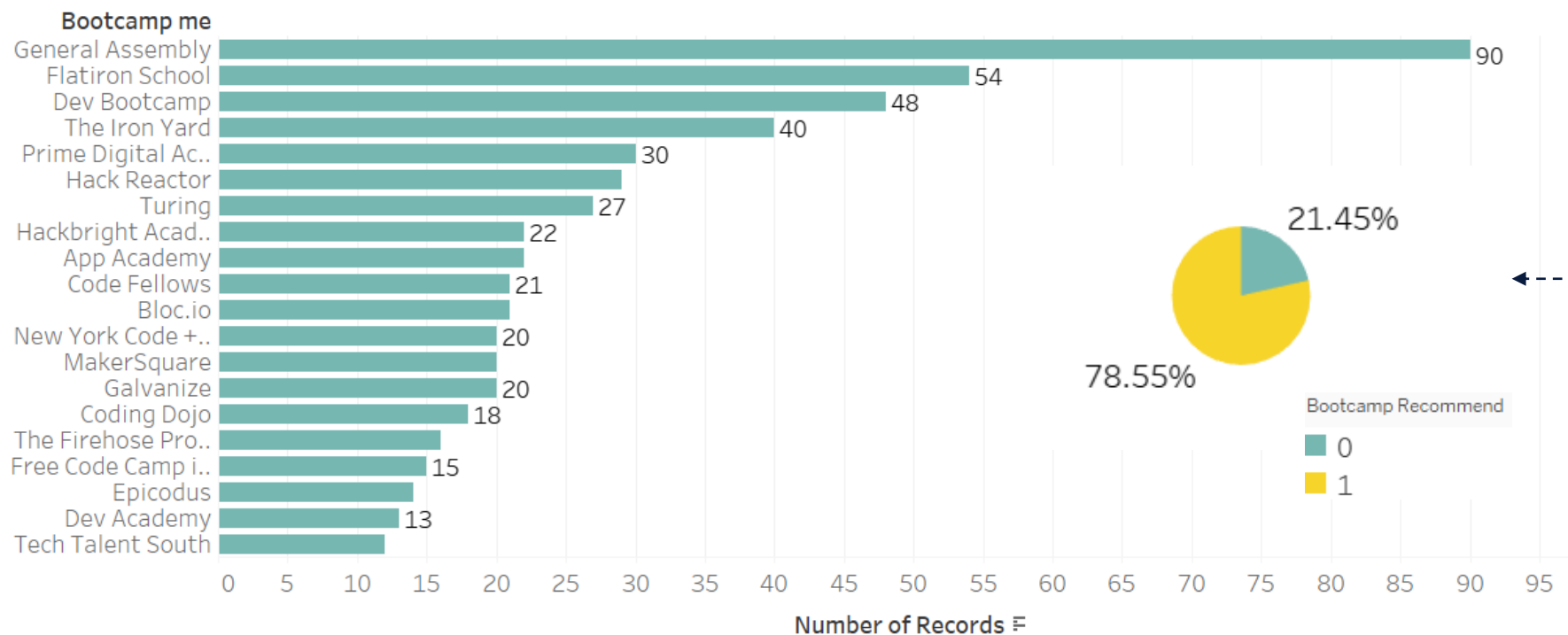| Podcast	Value
Sum of Podcast None	11,040
Sum of Podcast Code Newbie	1,878
Sum of Podcast JS Jabber	
Sum of Podcast SE Daily	775
Sum of Podcast Change Log	
Sum of Podcast Shop Talk	83
Sum of Podcast Ruby Rogues	
Sum of Podcast Developer Tea	80
Sum of Podcast Talk Python	
Sum of Podcast Programming Throw Down	34
Sum of Podcast Js Air	
Sum of Podcast Dot Net Rocks	18
Sum of Podcast Hanselminutes	
Sum of Podcast Coding Blocks	16
Sum of Podcast Web Ahead	

Value

1 market overview

1.1 competitive landscape

Only 6% of learners attend bootcamp. 68% of participants completed, of whom 50% found a job, the annual salary was concentrated in the 40k-80k dollars, and most of them were willing to recommend to others.

31.08%

68.92%

41.57%

58.43%

21.45%

78.55%

BootcampName

Bootcamp me

BootcampName	Number of Records
General Assembly	90
Flatiron School	54
Dev Bootcamp	48
The Iron Yard	40
Prime Digital Ac..	30
Hack Reactor	29
Turing	27
Hackbright Acad..	22
App Academy	22
Code Fellows	21
Bloc.io	21
New York Code +..	20
MakerSquare	20
Galvanize	20
Coding Dojo	18
The Firehose Pro..	16
Free Code Camp i..	15
Epicodus	14
Dev Academy	13
Tech Talent South	12

Number of Records

Bootcamp Recommend
- 0
- 1

Bootcamp Post Salary (bin)

Number of Records

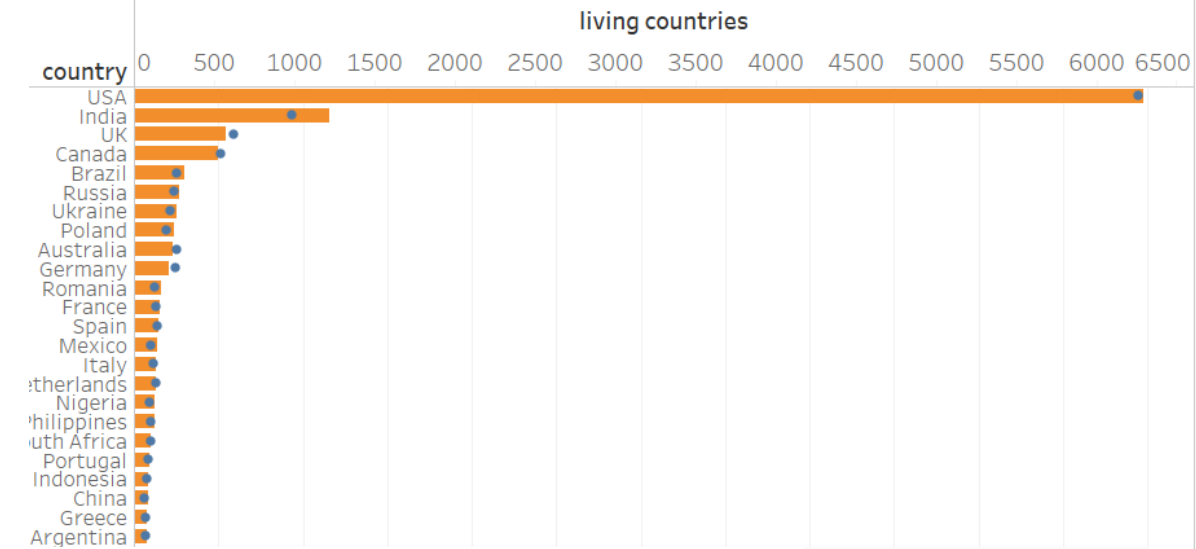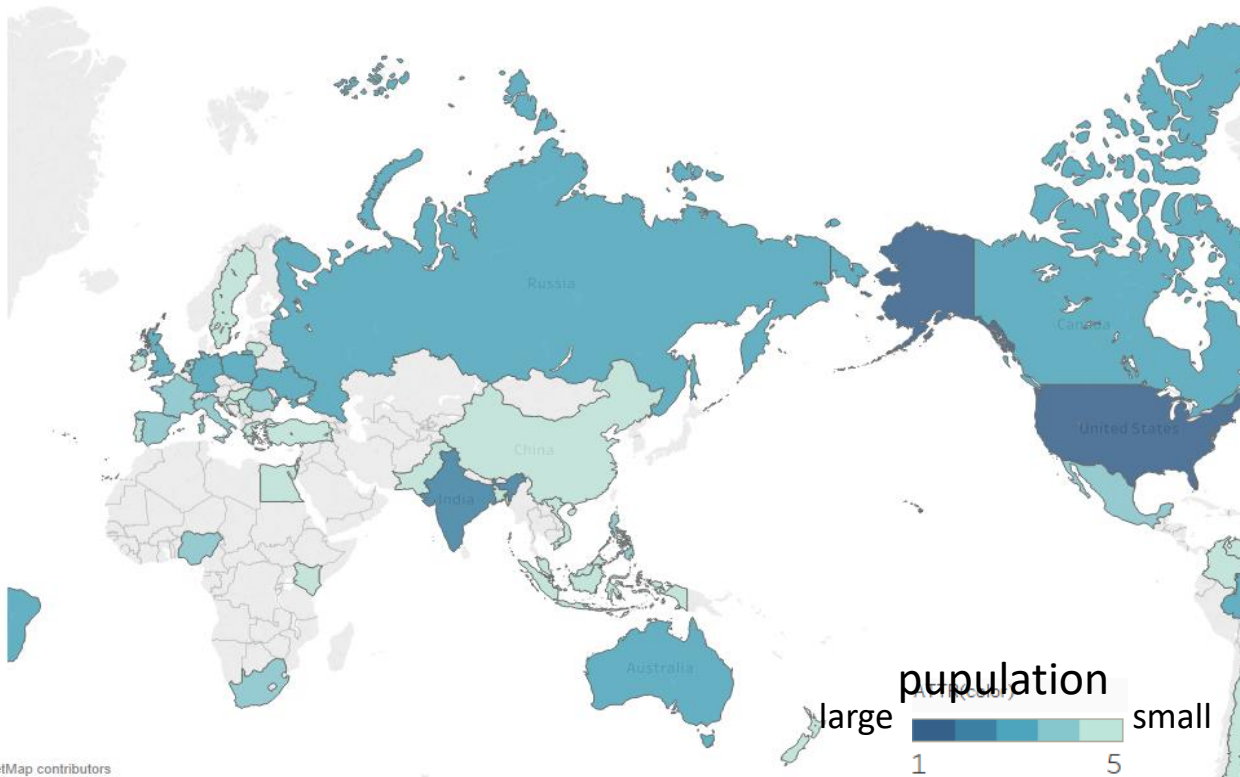$0K $20K $40K $60K $80K $100K $120K $140K
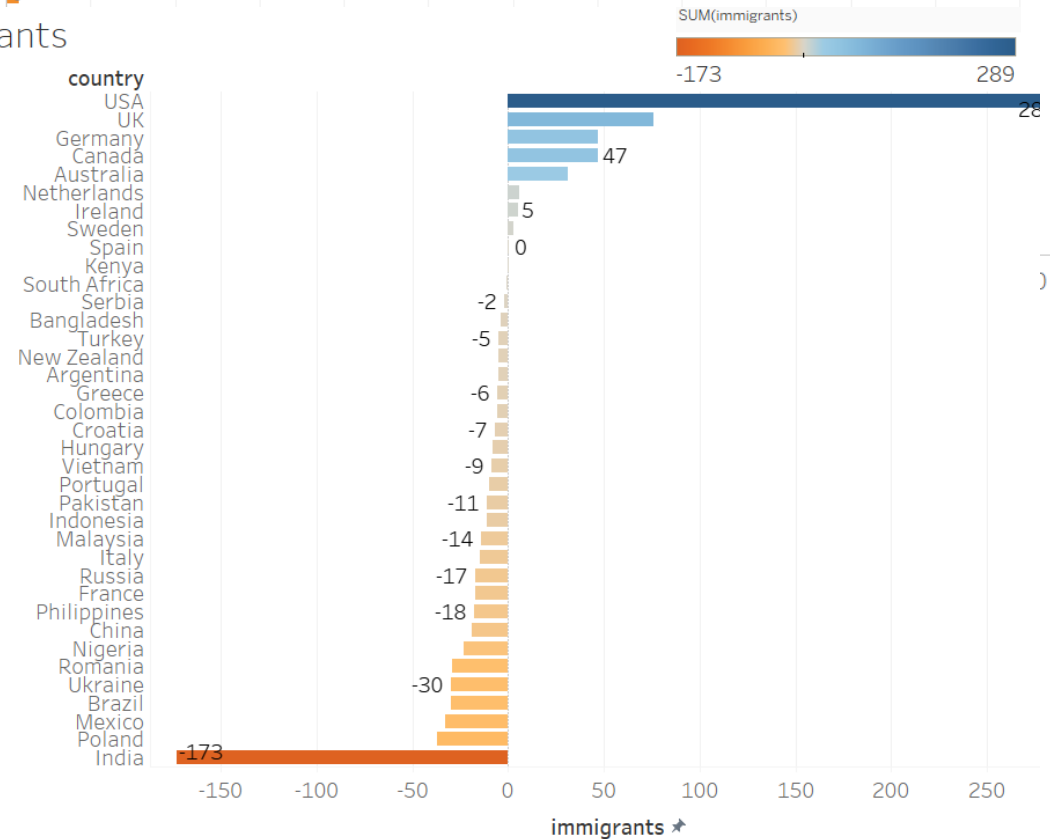
1 market overview

1.2 market analysis by region

Most customers are form USA and India. Besides, Immigration also plays a role, with Indian immigrants in the United States mostly.
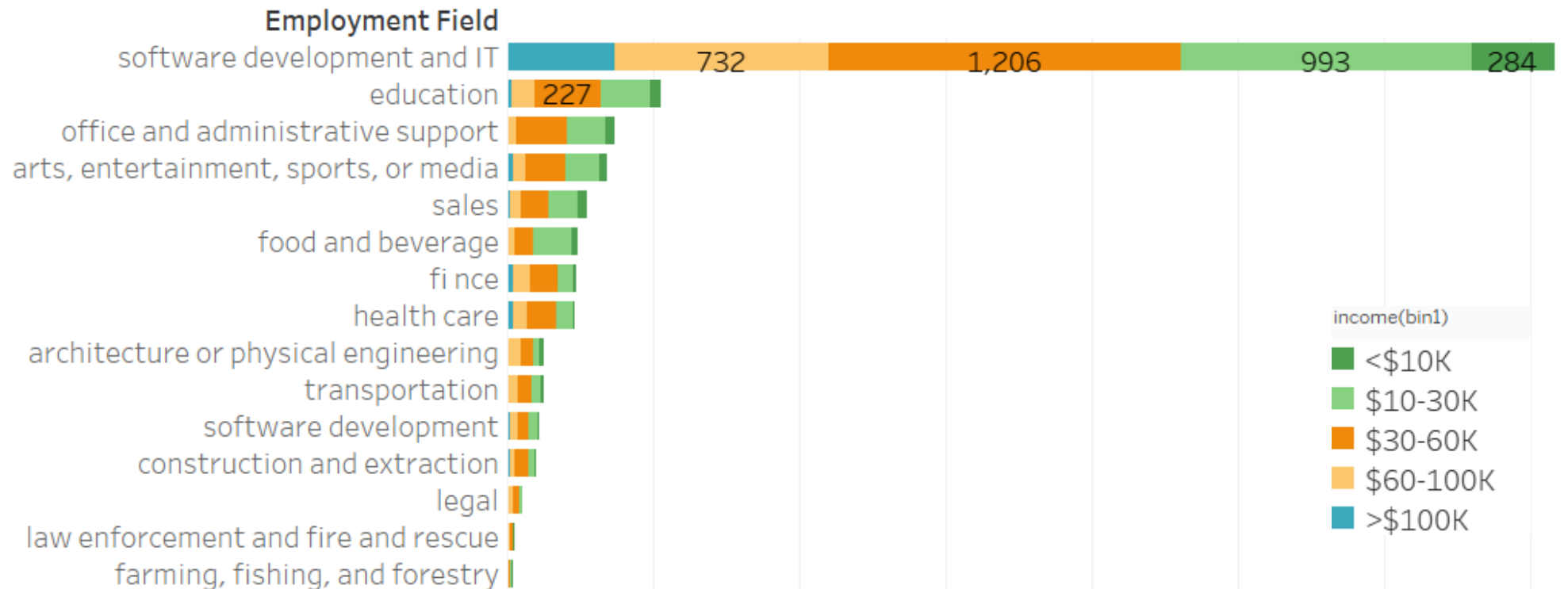

country citizens of learners

1 market overview

1.2 market analysis by industry fields

- Half are in a technology-related field. Others are from education, management, arts, etc.
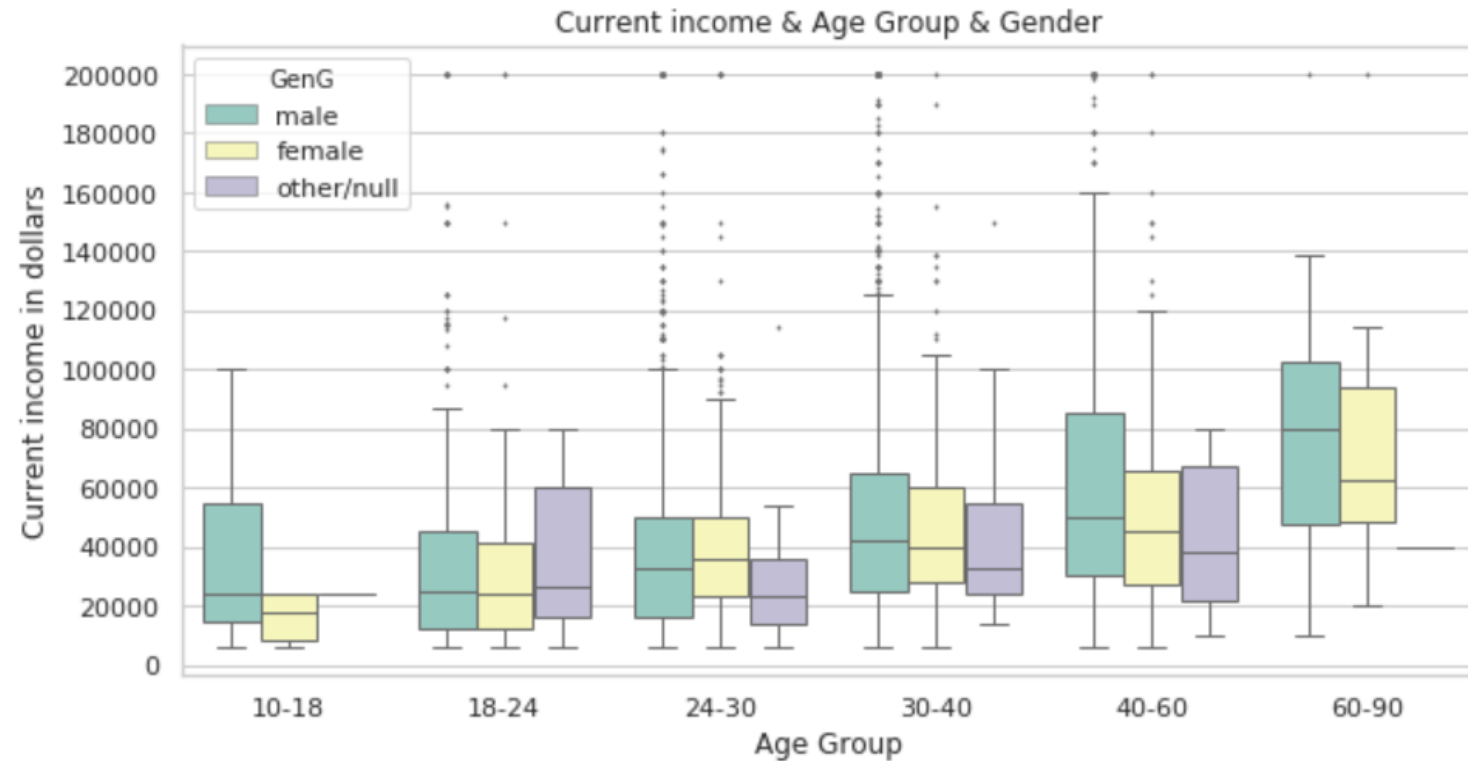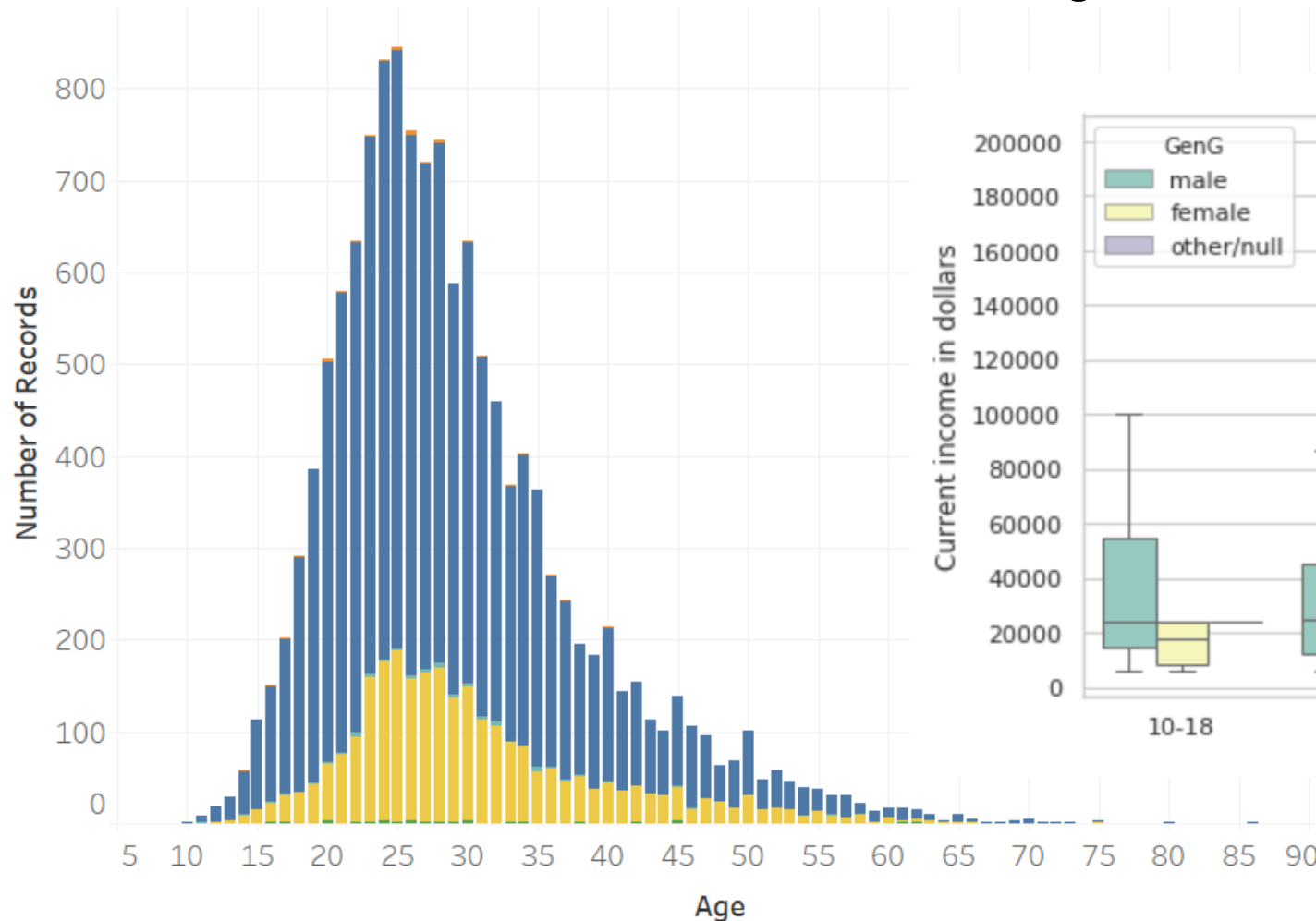- High income levels are also mainly in this field.

employfield

2.1 customer features

Most customers are aged between 20-35, and most are male. After the age of 30, men will earn more than women and other genders, and the gap will increase with age
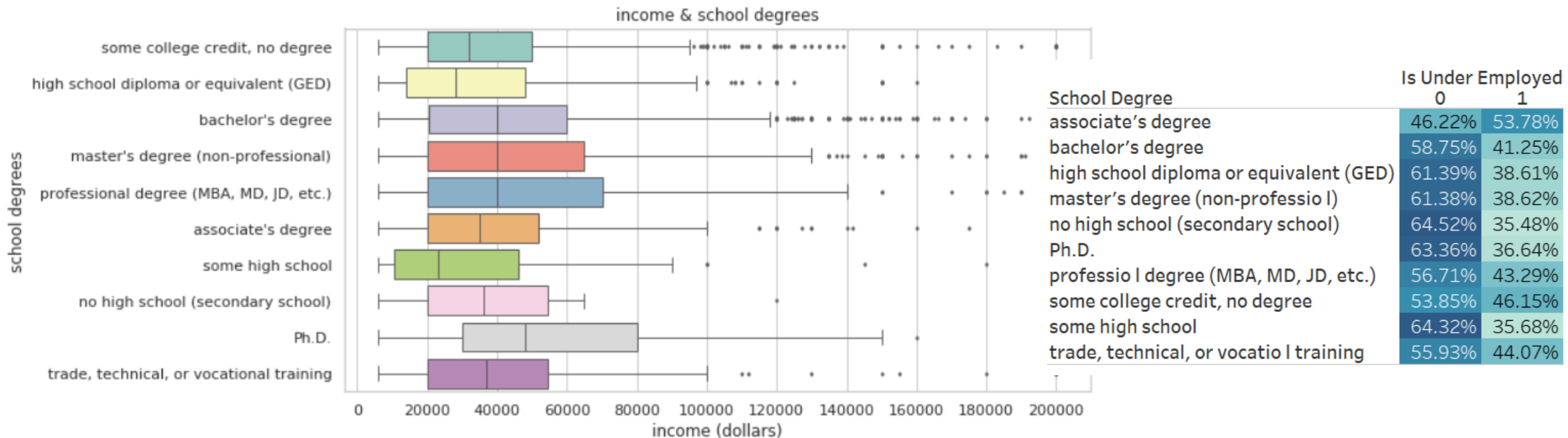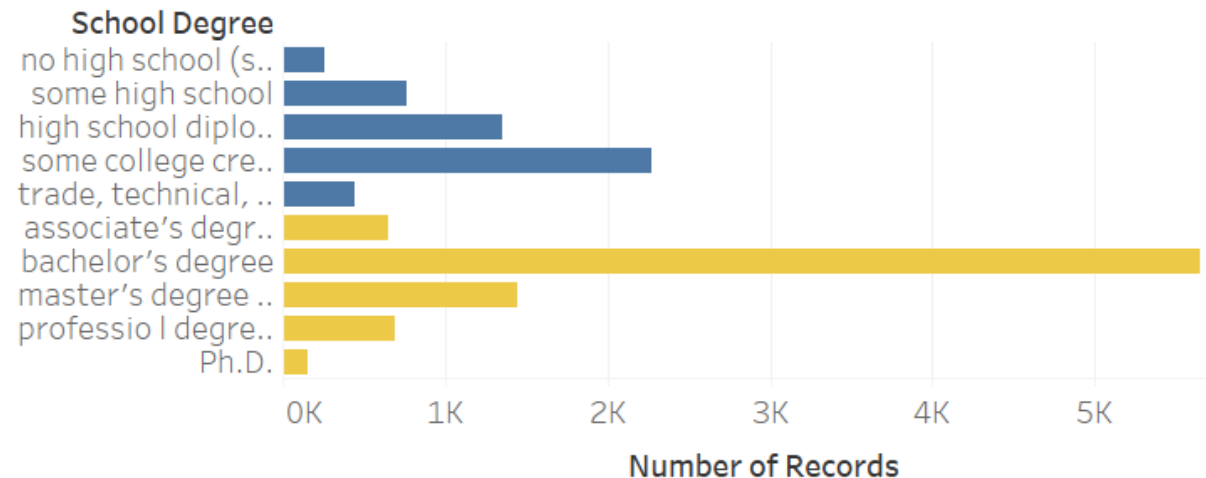
age&Gender

2.1 customer features

- Half of them earned at least a bachelor's degree.
- Higher education, higher annual income
- Non-traditional education like associate's degree, non-degree college education, tend to think they are underemployed.

School Degree

Number of Records

income & school degrees

| School Degree	Is Under Employed	
	0	1
associate's degree	46.22%	53.78%
bachelor's degree	58.75%	41.25%
high school diploma or equivalent (GED)	61.39%	38.61%
master's degree (non-professio l)	61.38%	38.62%
no high school (secondary school)	64.52%	35.48%
Ph.D.	63.36%	36.64%
professio l degree (MBA, MD, JD, etc.)	56.71%	43.29%
some college credit, no degree	53.85%	46.15%
some high school	64.32%	35.68%
trade, technical, or vocatio l training	55.93%	44.07%

2 customer features and behavior

2.2 customer behaviors

- Most start programming less than a year and learn less than 20 hours weekly.
- The longer it takes to program, the more likely he is a software developer and the higher the income
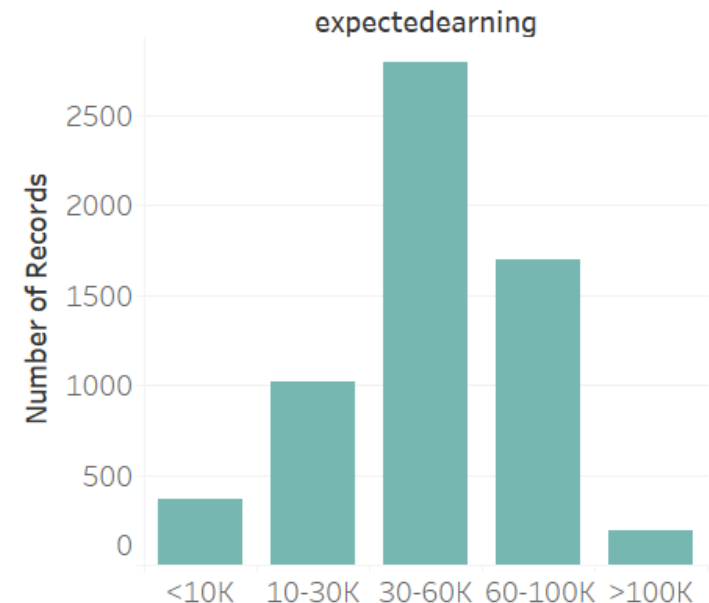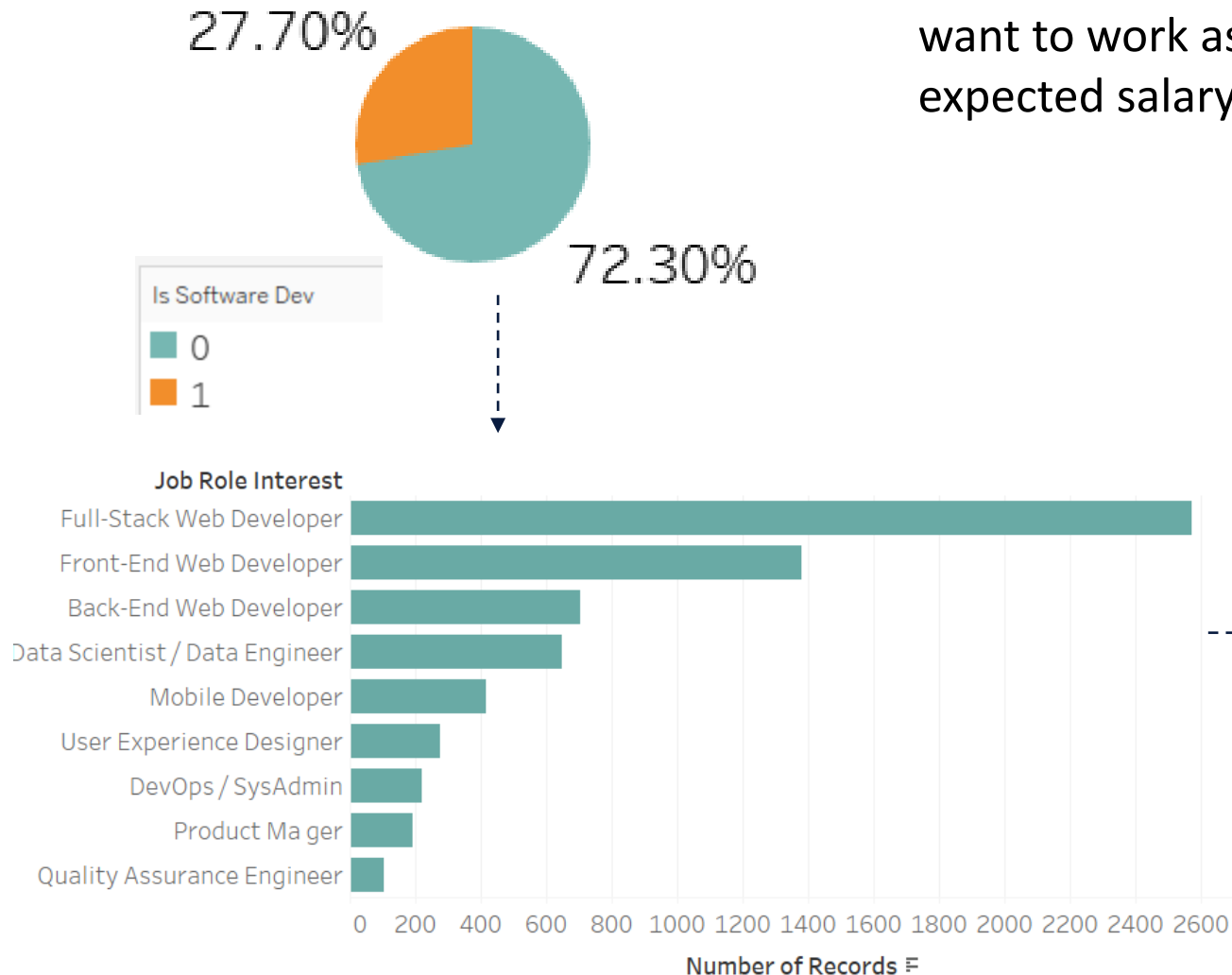

developer/month/hours
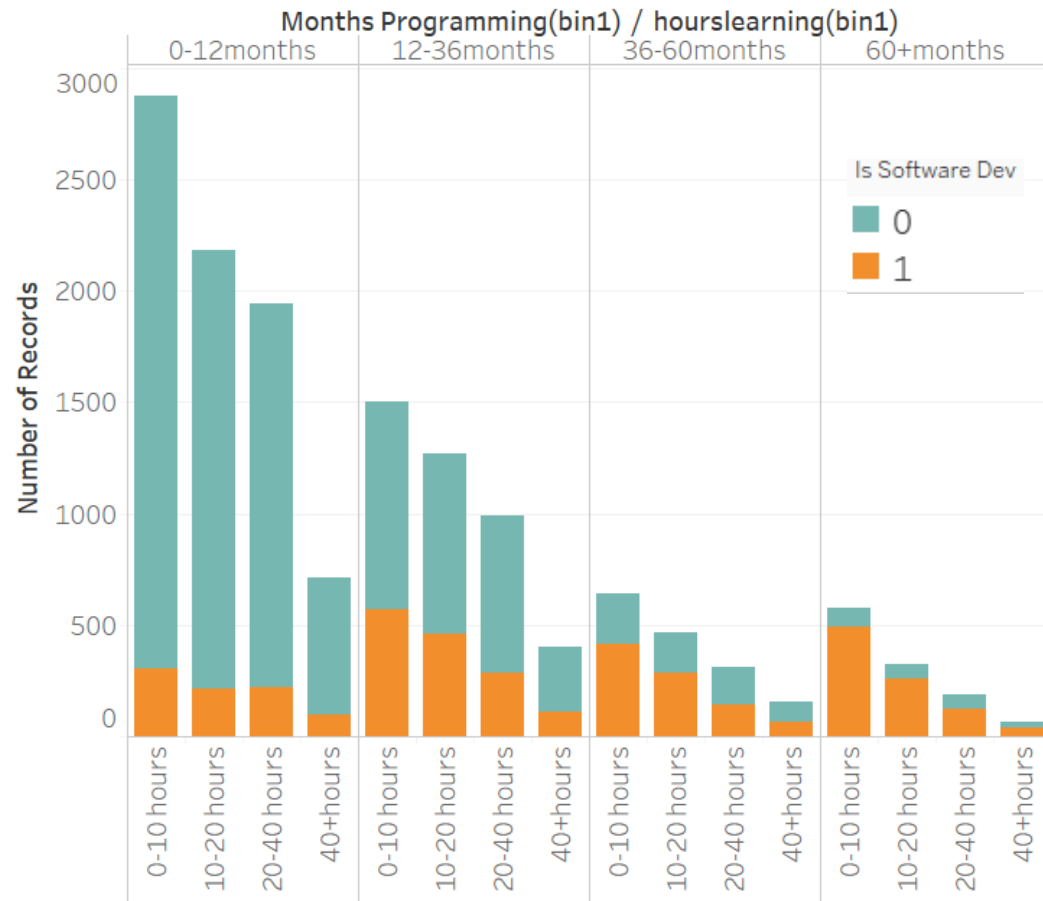
income/mon/hour
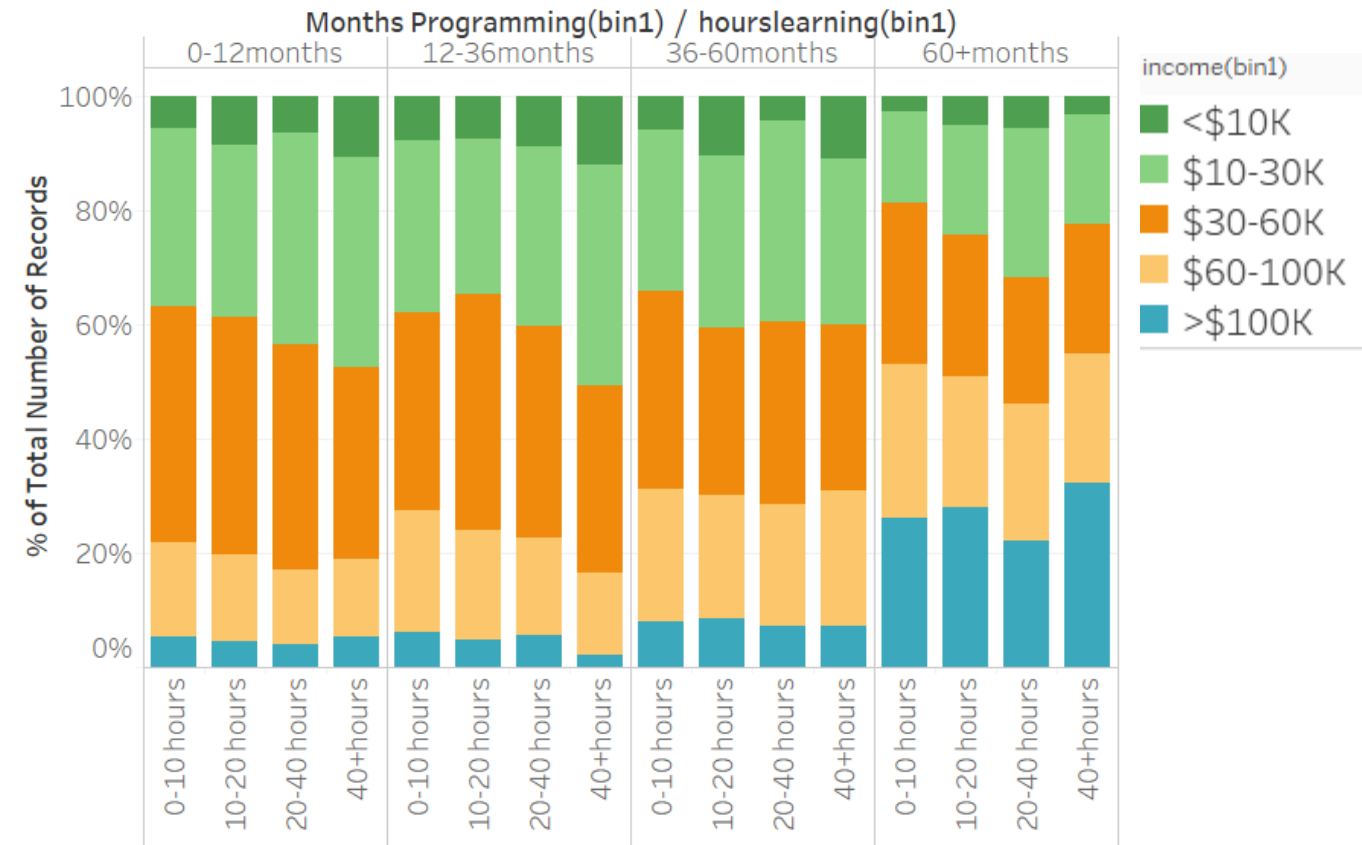
2 customer features and behavior

2.2 customer behaviors

- 72% learner are not software developers yet, but most of them want to work as a developer, especially web developers with expected salary of 30k-60k dollars per year.

ten classic personas (clustering)

- The customer are very diverse. To reveal the similarities or dissimilarities between customers, I did the clustering.

- Data preparation:
➢ turn the categories strings into number code
➢ remove the unrelated columns, such as ID
➢ normalize the data

- Since it is not clear what a good choice for the number k of clusters should be, I ran k-means for increasing values of k until the clustering error is below a prescribed level. Based on the plot, I chose 10 clusters.

- We can see most of the customers belong to cluster 2, 1,4 and 8.

The number of clusters vs clustering error
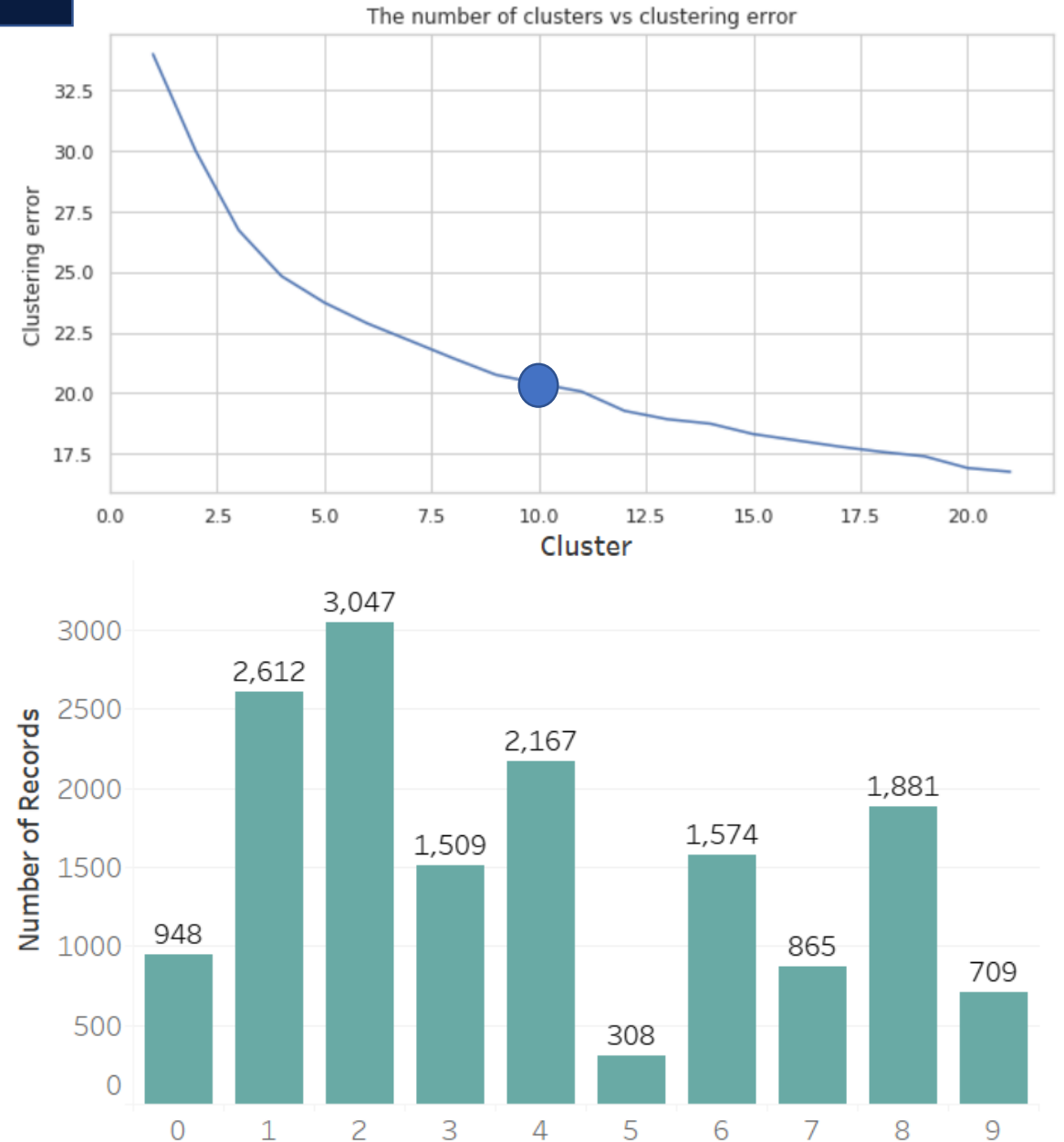
ten classic personas (clustering)

I explored the features of cluster 2 and 4, and find that

Cluster 2

I am not from technology field and not a software developer but want to transform to work as a web-developers in multinational or medium-size companies. I will apply the developer jobs within next 12 months.

I am already a software developer and have higher salary. I have the entrepreneur spirit and want to start my own business, work for a start-up companies or be freelance.

Cluster4

More analysis for other clusters, can be explored in tableau file

Thank You!