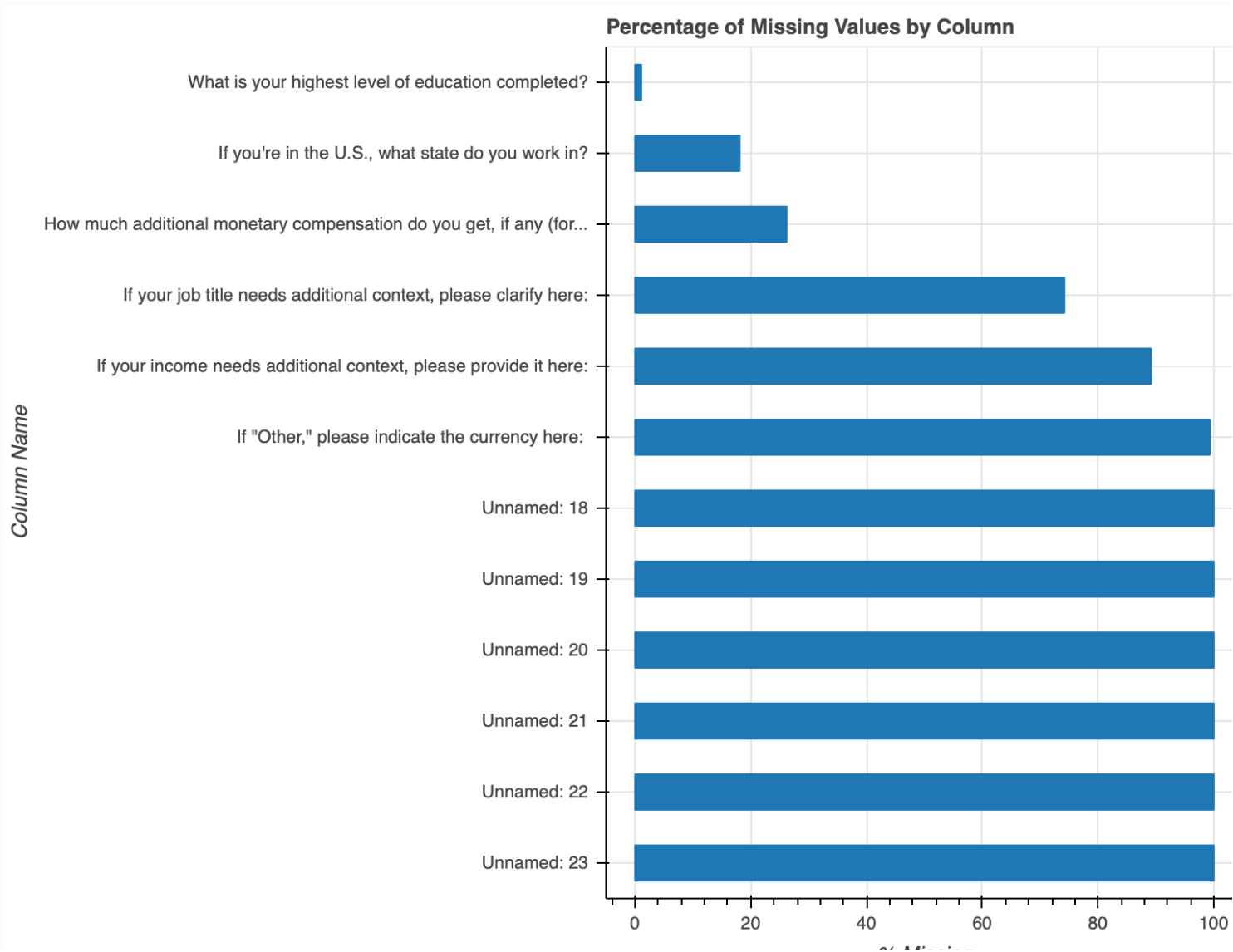# Data Profiling Report

Recommendations

- There are 24 columns with missing values in the dataset. Review missing values table below to assess severity.
- 92 rows are duplicated over all columns. There are 24 columns with duplicate values in the dataset. Review duplicates table below.
- There are 9 columns with free form text. Review free-form-text table below for handling decisions.
- There are 17 columns with datatype mismatch in the dataset. These columns have a datatype that is different from majority datatype of cell values. Review datatype mismatch table below.
- There are 18 columns with outliers in the dataset. For numeric columns this refers to numeric value, for non-numeric columns this refers to value length. Refer to misfielded values table for handling decision.
- There are 4 columns with potential misfielding/incorrect values in the dataset. Refer to misfielding tables below to verify these entries.

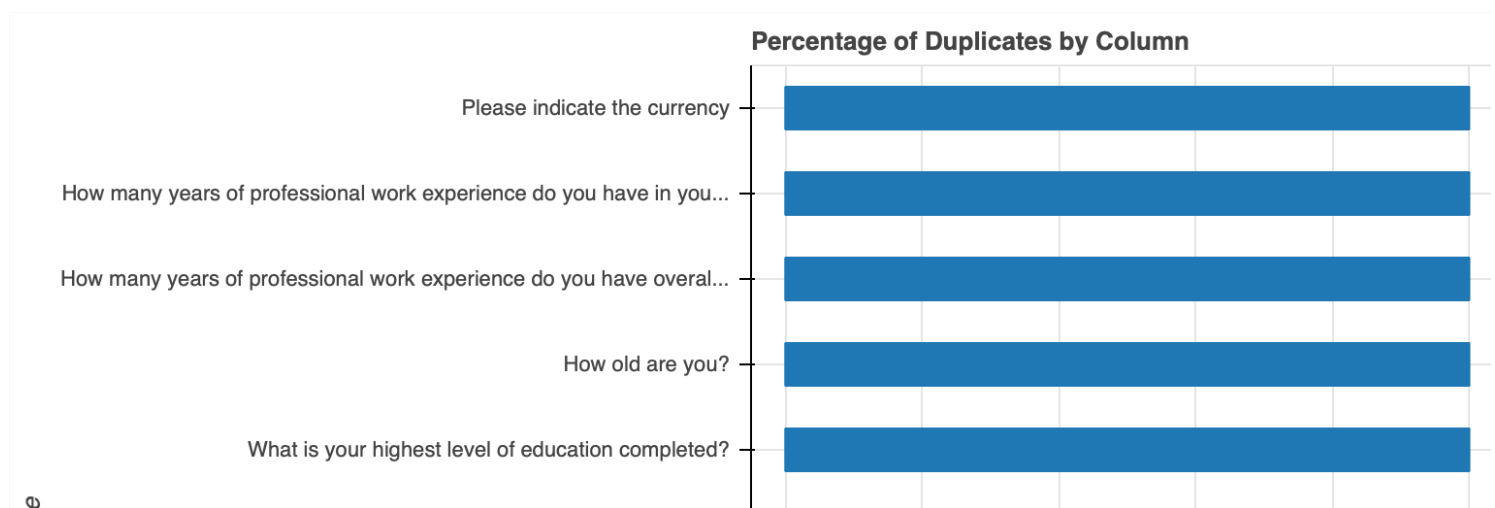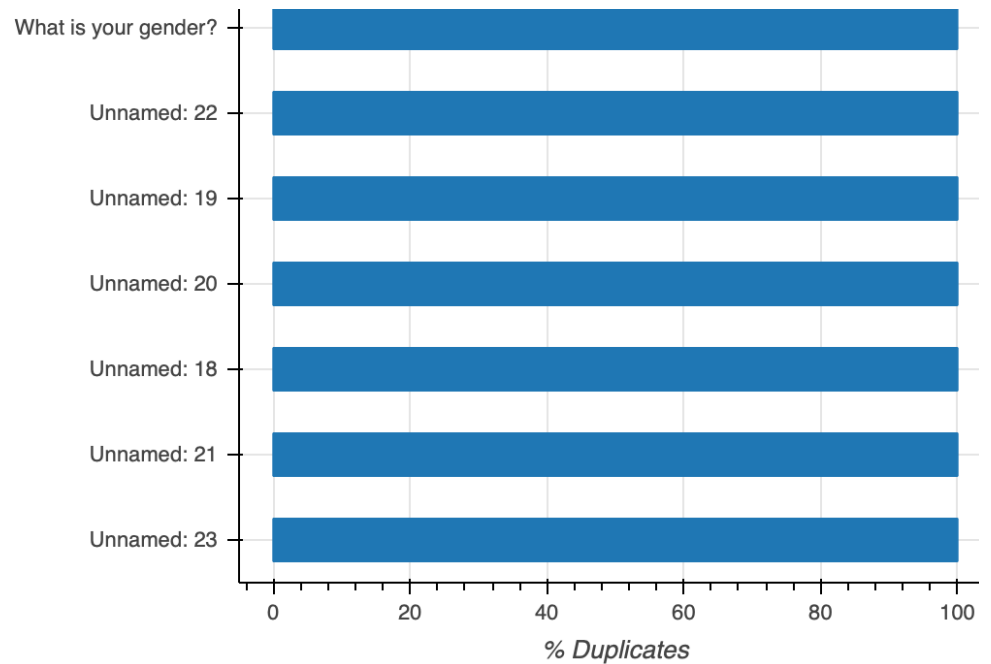### Anomolous Conditions Detected

Percentage of Missing Values

% Missing

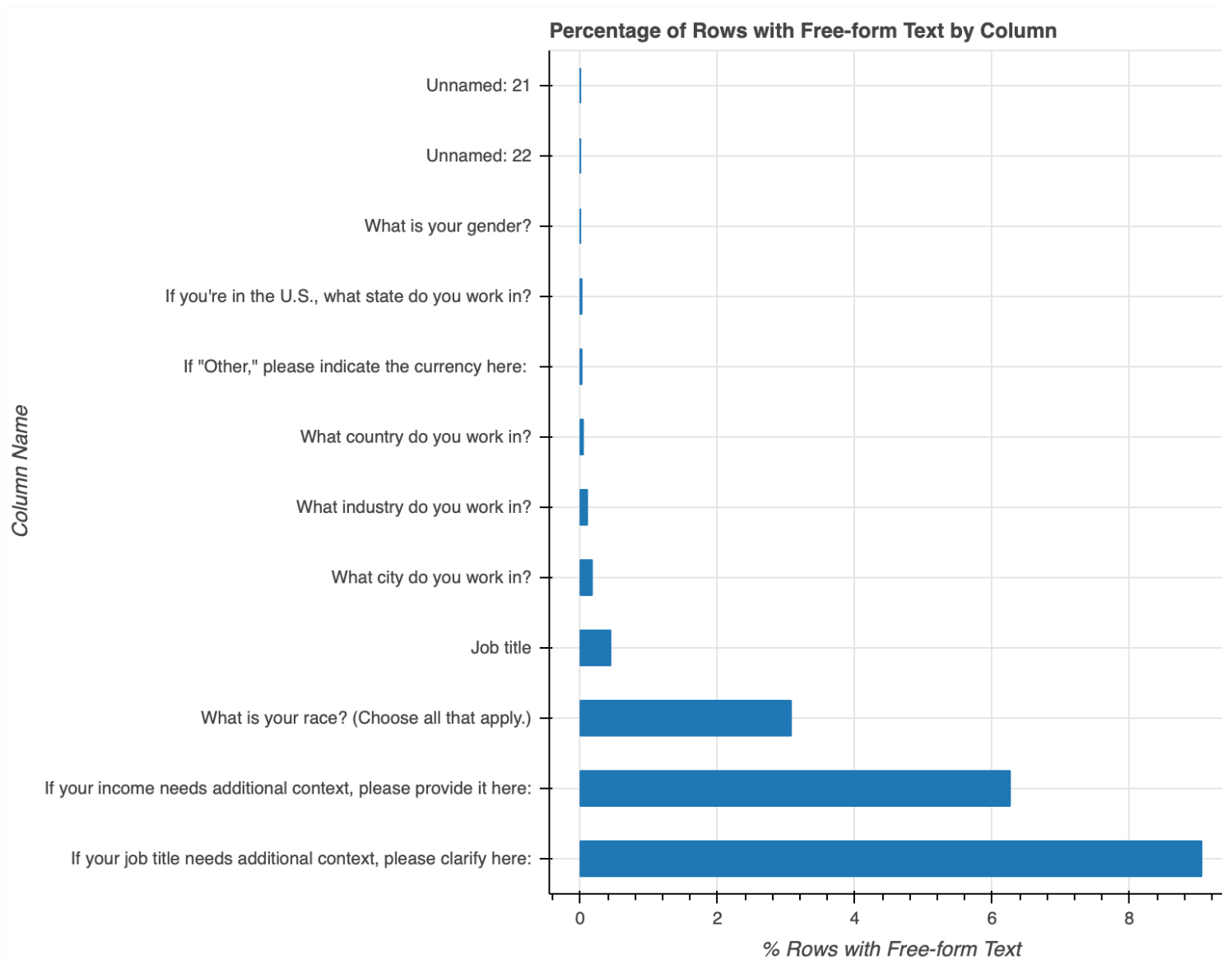|	Column Name	Missing %
23	Unnamed: 23	100.000000
22	Unnamed: 22	100.000000
21	Unnamed: 21	100.000000
20	Unnamed: 20	100.000000
19	Unnamed: 19	100.000000
18	Unnamed: 18	100.000000
8	If "Other," please indicate the currency here:	99.297207
9	If your income needs additional context, please provide it here:	89.176269
4	If your job title needs additional context, please clarify here:	74.199993
6	How much additional monetary compensation do you get, if any (for...	26.203132
11	If you're in the U.S., what state do you work in?	18.097820
15	What is your highest level of education completed?	1.091649
17	What is your race? (Choose all that apply.)	0.931112
16	What is your gender?	0.923977
12	What city do you work in?	0.599336
2	What industry do you work in?	0.588634
14	How many years of professional work experience do you have in you...	0.331776
13	How many years of professional work experience do you have overal...	0.331776
1	How old are you?	0.331776
10	What country do you work in?	0.331776
7	Please indicate the currency	0.331776
5	What is your annual salary? (You'll indicate the currency in a la...	0.331776
3	Job title	0.331776
0	Timestamp	0.331776

Duplicates

Provided dataset has 92 duplicate rows.

Percentage of Duplicates by Column

Column Nan

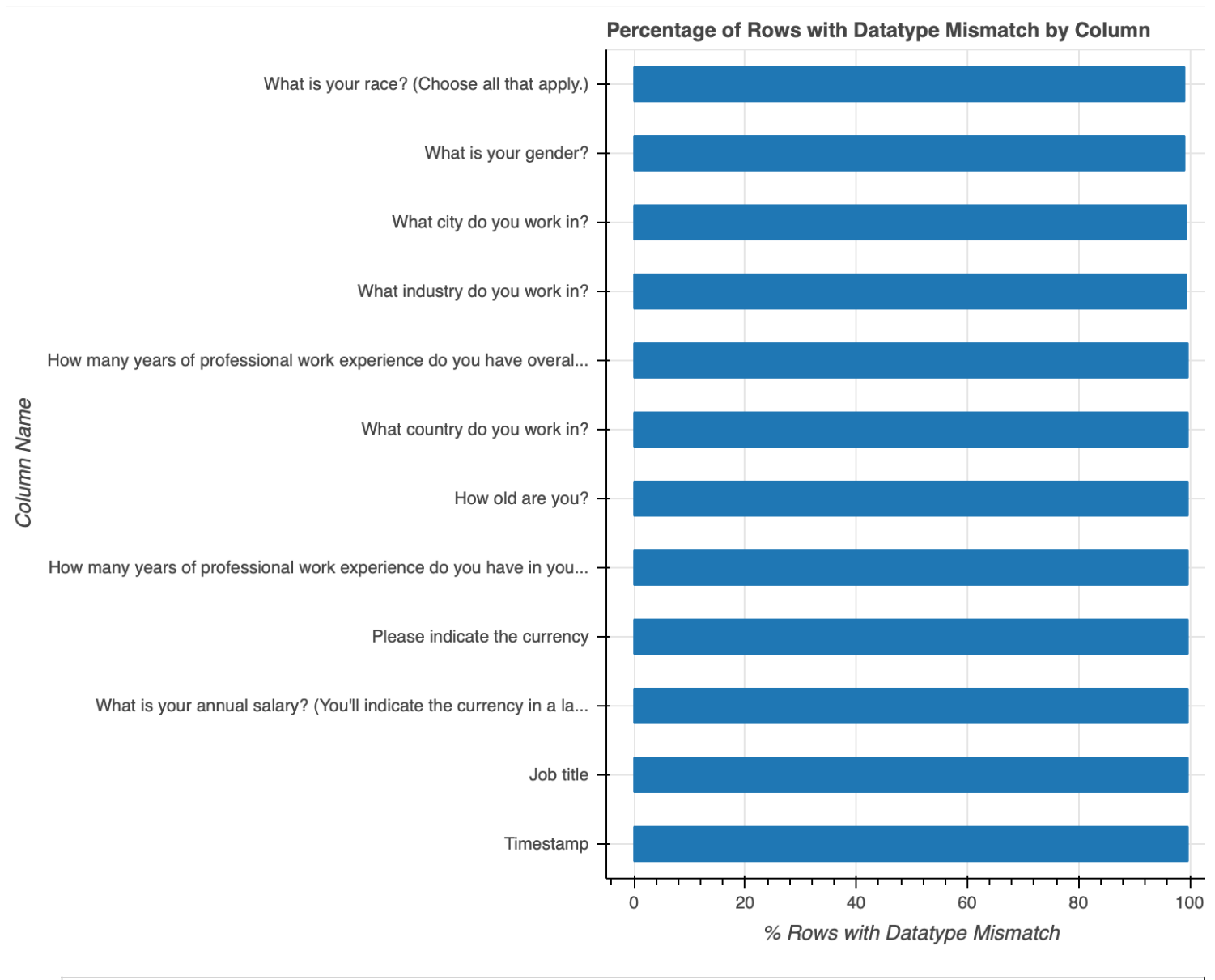|	Column Name	Duplicate Percentage
23	Unnamed: 23	99.996433
21	Unnamed: 21	99.996433
18	Unnamed: 18	99.996433
20	Unnamed: 20	99.996433
19	Unnamed: 19	99.996433
22	Unnamed: 22	99.996433
16	What is your gender?	99.978595
15	What is your highest level of education completed?	99.975028
1	How old are you?	99.971460
13	How many years of professional work experience do you have overal...	99.967893
14	How many years of professional work experience do you have in you...	99.967893
7	Please indicate the currency	99.957190
17	What is your race? (Choose all that apply.)	99.825194
8	If "Other," please indicate the currency here:	99.589740
11	If you're in the U.S., what state do you work in?	99.529093
10	What country do you work in?	98.683600
6	How much additional monetary compensation do you get, if any (for...	96.989048
2	What industry do you work in?	95.683351
9	If your income needs additional context, please provide it here:	89.397453
5	What is your annual salary? (You'll indicate the currency in a la...	84.713353
12	What city do you work in?	82.883236
4	If your job title needs additional context, please clarify here:	75.127537
3	Job title	49.024295
0	Timestamp	10.256502

Free Form Text in Dataset

Percentage of Rows with Free-form Text by Column

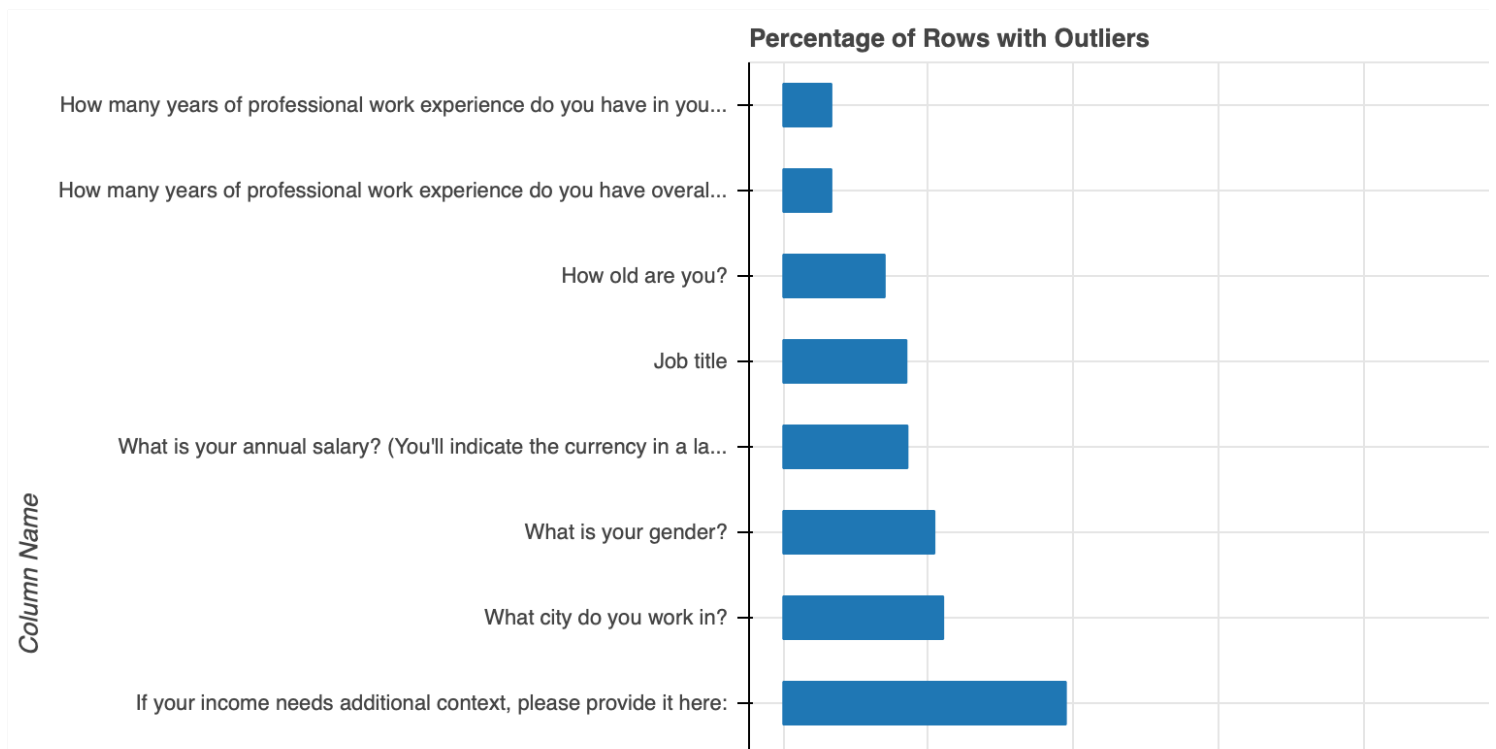|	Column Name	Percentage of Rows with Free-form Text
4	If your job title needs additional context, please clarify here:	9.05
9	If your income needs additional context, please provide it here:	6.26
17	What is your race? (Choose all that apply.)	3.07
3	Job title	0.44
12	What city do you work in?	0.17
2	What industry do you work in?	0.10
10	What country do you work in?	0.04
8	If "Other," please indicate the currency here:	0.02
11	If you're in the U.S., what state do you work in?	0.02
16	What is your gender?	0.00
22	Unnamed: 22	0.00
21	Unnamed: 21	0.00
20	Unnamed: 20	0.00

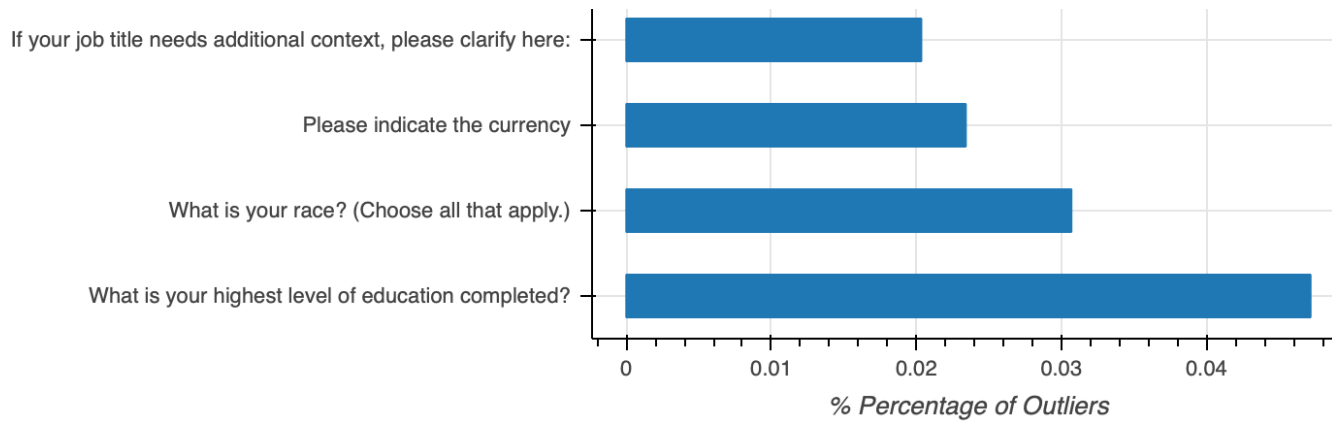| 19	Unnamed: 19	0.00
18	Unnamed: 18	0.00
0	Timestamp	0.00
15	What is your highest level of education completed?	0.00
14	How many years of professional work experience do you have in you...	0.00
13	How many years of professional work experience do you have overal...	0.00
1	How old are you?	0.00
7	Please indicate the currency	0.00
6	How much additional monetary compensation do you get, if any (for...	0.00
5	What is your annual salary? (You'll indicate the currency in a la...	0.00
23	Unnamed: 23	0.00

Datatype Mismatch in Dataset

Percentage of Rows with Datatype Mismatch by Column

	Column Name	Column Datatype	Recommended Datatype	Percentage of Recommended Cells
0	Timestamp	object	str	99.67
3	Job title	object	str	99.67
5	What is your annual salary? (You'll indicate the currency in a la...	object	str	99.67
6	Please indicate the currency	object	str	99.67
13	How many years of professional work experience do you have in you...	object	str	99.67
1	How old are you?	object	str	99.67
9	What country do you work in?	object	str	99.67
12	How many years of professional work experience do you have overal...	object	str	99.67
2	What industry do you work in?	object	str	99.41
11	What city do you work in?	object	str	99.40
15	What is your gender?	object	str	99.08
16	What is your race? (Choose all that apply.)	object	str	99.07
14	What is your highest level of education completed?	object	str	98.91
10	If you're in the U.S., what state do you work in?	object	str	81.90
4	If your job title needs additional context, please clarify here:	object	str	25.80
8	If your income needs additional context, please provide it here:	object	str	10.82
7	If "Other," please indicate the currency here:	object	str	0.70

Percentage of Rows With Outliers

Percentage of Rows with Outliers

	Column Name	Percentage of Outliers
15	What is your highest level of education completed?	0.047162
17	What is your race? (Choose all that apply.)	0.030680
7	Please indicate the currency	0.023403
4	If your job title needs additional context, please clarify here:	0.020335
9	If your income needs additional context, please provide it here:	0.019478
12	What city do you work in?	0.011024
16	What is your gender?	0.010417
5	What is your annual salary? (You'll indicate the currency in a la...	0.008562
3	Job title	0.008491
1	How old are you?	0.006992
13	How many years of professional work experience do you have overal...	0.003318
14	How many years of professional work experience do you have in you...	0.003318
0	Timestamp	0.003318
10	What country do you work in?	0.002604
8	If "Other," please indicate the currency here:	0.001891
11	If you're in the U.S., what state do you work in?	0.001748
2	What industry do you work in?	0.001249
6	How much additional monetary compensation do you get, if any (for...	0.000036

Potentially Misfielded Values

There are 11 columns with potentially misfielded values:

	Timestamp
0	5/1/2021 0:01:23
1	5/1/2021 0:03:47
2	5/1/2021 0:05:57
3	5/1/2021 0:07:13

	What industry do you work in?
0	Library (its a non-profit and its a govt job - how would I list that? Not all libraries are govt jobs)
1	Public Library (might be considered Government, but that always seems an odd designation...)
	Language Services company, unsure the broad category to use. Our clients are branding agencies, and their

2	clients are frequently (but not always) pharmaceutical companies.
3	Hybrid Nonprofit Higher Education (we are part of a university but our entire budget comes from grants)

	Job title
0	Solutions Control Specialist (Chemist)/Wastewater Pretreatment Facility Chief Operator
1	Associate Vice President of Advancement Administration & Chief Annual Giving Officer
2	No official title. Do whatever needs to be done- i was hired as a bookeeper, now do estimatinf, projext management, schedulong
3	NaN

	If your job title needs additional context, please clarify here:
0	Surface level, I perform the basic duties of a senior-level sales associate in a small boutique. Due to COVID, along with the owner's tragic and early passing in August of 2020 (she had been in and out of the store for two years and eventually the breast cancer killed her) our now bare-bones team delegates all the extra tasks. I am considered to be a marketing coordinator for the team, but I want to preface that with the explanation that it's not a traditional version of the role, nor is it paid as such.
1	Senior Consultant is the standard title for newly minted MBAs or holders of other graduate degrees at my firm. You can become an SC by starting as an analyst and working your way up for 4-5 years, but most people either come in with an MBA or leave to get one and then come back, because it's hard to get to manager level without one. As an SC, I am responsible for managing a workstream (usually yourself and a few analysts) as part of a larger project. I don't do any business development.
2	I help students with internships, job shadowing, plan structured field trips. I also help students with their after high school plans and provide some services on campus or connect them to services. I do career assessments and interpretations, un-muddle job applications (why, why, WHY must you complete both a resume AND fill out the job experience? And WHY are you asking for resumes for part-time, entry-level, seasonal work advertised for high school students?). I'm part of the School Counseling team (NOT GUIDANCE COUNSELING!!!!) and my "other duties as assigned" typically involves some graphic design, subbing, social/emotional support (not services- I'm a friendly face not a counselor), testing, event planning, college essay proof-reader, and, in a pinch, IT support.
3	The role is fairly nebulous because it fits the needs of the unit or department. Essentially I am the assistant manager with all functions of a manager, but don't have authority over hiring decisions, termination decisions, and can't do significant disciplinary action where termination is a result. In my role I manage team member scheduling, attendance, quality and safety reporting, I create education for the team, I manage compliance with education, I also manage quality improvement and evidence based practice projects, I also teach classes for our team. This is all with the expectation that I work as a charge nurse for a set amount of my hours. For some they have a 2/3rds of their time is working floor and 1/3rd indirect. This is dependant on the manager of the unit and their needs for the role.

	What is your annual salary? (You'll indicate the currency in a later question. If you are part-time or hourly, please enter an annualized equivalent -- what you would earn if you worked the job 40 hours a week, 52 weeks a year.)
0	58
1	35
2	38
3	61

	If "Other," please indicate the currency here:
0	I work for an online state university, managing admissions data. Not direct tech support.
1	AUD and NZD aren't the same currency, and have absolutely nothing to do with each other :(
2	Many non-salary benefits - travel, free healthcare for self, very low for family, non-taxable housing allowance
3	additional compensation is for overtime (i am paid hourly) so it varies. i have included an estimate

	If your income needs additional context, please provide it here:
0	Actual work is 60+ hours weekly (5x 12 hour shifts per week), over 3000 hours annually, not the 2080 annual hours for a 40-hour job. Paid entirely on commission (38% of fare after flag drop + tips, typically 10-20% of fare) and bonuses from associated businesses (liquor stores, clubs, massage places, etc. all pay drivers to bring passengers to them) of approximately $9000/year. \n\n1 week paid vacation after 1 year with the company. 1 week per year after that up to 4 weeks. A minority of drivers stay in the job that long. Barely affordable health insurance, none before ACA.\n\nThat was my last year of income, 2019. My income had fallen more than 25% per year for last 3 years thanks to rideshare services dumping their service into the market for below-cost prices.\n\nI have been unemployed since COVID shut down tourism. All of these answers are for 2019, the last full year I was employed.
1	I work part-time, about 0.5 FTE (20 hrs/wk). When I took the role, it was offered that I could also work 0.5 FTE in another office to get up to 1.0 FTE, but I declined. My boss has repeatedly offered me 1.0 FTE in my office, and I have repeatedly declined, as I am also taking university classes part time and because I would find this job too stressful at 1.0 FTE. I also work the minimum number of hours required to access health benefits, paid vacation, pension, etc. I get 3 weeks of paid vacation each year, plus an additional ~2 weeks off at Christmas. I frequently work a bit over my set hours each week, but we are told not to record the time (I know!) and just take it off as lieu time later, which I try to keep on top of. There is some expectation that I am "on call" and able to respond to texts in off hours, but I make sure to record any time that I spend on this outside of my regular hours.
2	My bonus pay is an essential part of my remuneration, and not optional. We must earn minimum 100% of our allotted incentive quota in each financial quarter, and this amount can range up to 200%, depending on your performance, and role. \n\nSo as a rule, I would definitely be earning at least Rs.90,00 (Approx.) every quarter, with this amount going up to Rs.180,000 depending on my performance. \n\nThere are clear numerical performance parameters set in place to measure 100% performance from each employee, so there's no negotiation involved. You know what you'll get, based on your pre-set KPIs. \n\nIt's an unusual system of compensation and makes it difficult to explain to other employers that my actual, set income is in fact Rs.1,800,000 and not Rs.1,400,000.\n\nIn fact, I was planning to write in to you with a letter later this week, because I've been wondering how to explain this to potential employers without sounding like a scammer! In India the practice of asking for previous salary and basing offers on that is still the norm, so the usual advice of "don't disclose your previous income" does not apply to my situation, I'm afraid.
3	Also provided with my housing, utilities, cable and internet. Pays for all my gas and a lot of repair on my vehicle which I use some for work. As it is a horse farm, when there is a horse I want and can ride I am able to do that. I am able to use the farm truck and horse trailer to haul to go places to ride. The farm also owns the majority of the tack that I use when riding. I am allowed pets. Because my employer also owns another company I have health insurance that they pay a small portion of and a retirement account that they match up to 3%. Some may not sound like a big deal in the corporate environment but for the horse business itâ€™s not bad. Some of my bonus money is based on horses successfully having babies and then how many horses I get pregnant for the next year. This can really fluctuate depending on not only Mother Nature but the goals of the farm. In years past we had 15-20 babies a year. Right now we are at 4-6.

	What country do you work in?
0	We don't get raises, we get quarterly bonuses, but they periodically asses income in the area you work, so I got a raise because a 3rd party assessment showed I was paid too little for the area we were located
1	United States (I work from home and my clients are all over the US/Canada/PR
2	I earn commission on sales. If I meet quota, I'm guaranteed another 16k min. Last year i earned an additional 27k. It's not uncommon for people in my space to earn 100k+ after commission.
3	I was brought in on this salary to help with the EHR and very quickly was promoted to current position but compensation was not altered.

	If you're in the U.S., what state do you work in?
0	Arkansas, Idaho, Kansas, Louisiana, Michigan, Mississippi, Nevada, New York, South Carolina, Tennessee, Washington
	Arkansas, Idaho, Kansas, Louisiana, Michigan, Mississippi, Nevada, New York, South Carolina, Tennessee,

1	Washington
2	California, District of Columbia, Illinois, Iowa, Maryland, Minnesota
3	California, Illinois, Massachusetts, North Carolina, South Carolina, Virginia

	What city do you work in?
0	Work remotely, but the company is based out of Chicago, IL and I work in Newark, NJ
1	In various cities. I work on projects through my employer for various financial organisations in the field of compliance & risk. My new project will be based in Amsterdam.
2	NaN
3	NaN

	What is your race? (Choose all that apply.)
0	Hispanic, Latino, or Spanish origin, Native American or Alaska Native, Another option not listed here or prefer not to answer
1	Asian or Asian American, Hispanic, Latino, or Spanish origin, Another option not listed here or prefer not to answer
2	Asian or Asian American, Hispanic, Latino, or Spanish origin, White, Another option not listed here or prefer not to answer
3	NaN