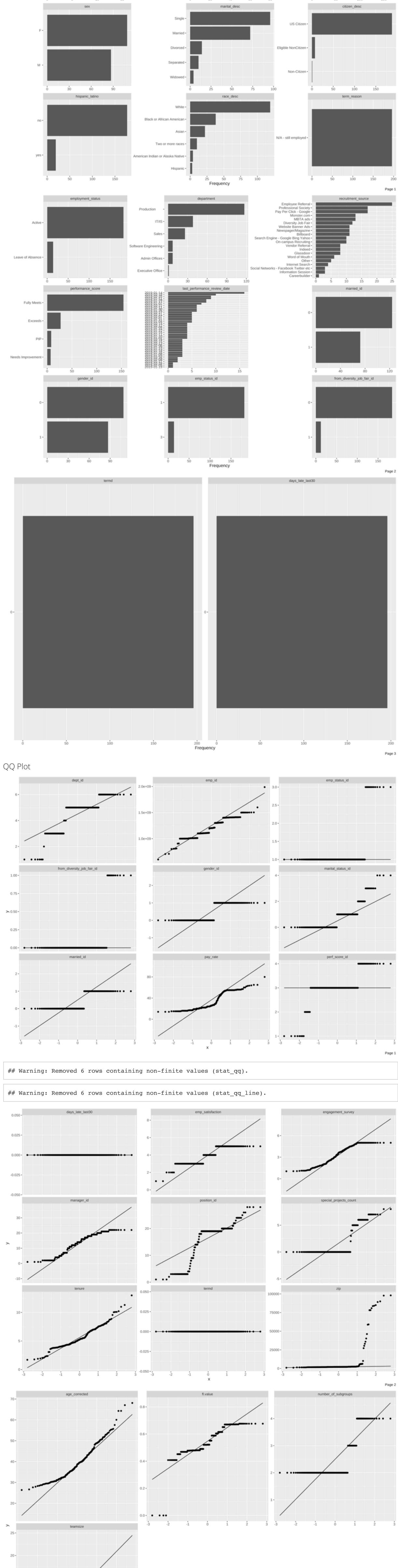
Data Profiling Report Basic Statistics Raw Counts Percentages Data Structure Missing Data Profile Univariate Distribution Histogram Bar Chart (with frequency) QQ Plot Correlation Analysis Principal Component Analysis Basic Statistics Raw Counts Value Name Rows 196 40 Columns Discrete columns 17 Continuous columns 22 All missing columns 1 Missing observations202 Complete Rows 0 Total observations 7,840 Memory allocation 97.7 Kb Percentages Memory Usage: 97.7 Kb 42.500% Discrete Columns -55.000% Continuous Columns -Dimension column 2.500% All Missing Columns observation Complete Rows - 0.000% Missing Observations -Value Data Structure omanager_name (chr) oemployee name (chr) oemployee name (chr) oemp_id (int) omarried_id (int) omarried_status_id (int) oemp_status_id (int) oemp_status_id (int) oemp_status_id (int) oemp_status_id (int) operf_score_id (int) operf_score_id (int) operf_score_id (int) opay_rate (num) otermd (int) oposition_id (int) oposition_id (int) oposition_id (int) ostate (chr) ozin (int) odob (chr) osex (chr) omarital_desc (chr) ocitizen_desc (chr) orace_desc (chr) odateof_hire (chr) odateof_termination (logi) oterm_reason (chr) oemployment_status (chr) odepartment (chr) odepartment (chr) omanager_id (int) orecruitment_source (chr) operformance_score (chr) oengagement_survey (num) oemp_satisfaction (int) ospecial_projects_count (int) olast_performance_review_date (chr) odays_late_last30 (int) otenure (num) oage_corrected (num) oteamsize (int) onumber_of_subgroups (int) off.value (num) root (Classes 'data.table' and 'data.frame': 196 obs. of 40 variables:) Missing Data Profile fl.value number_of_subgroups teamsize age_corrected tenure days_late_last30 last_performance_review_date special_projects_count emp_satisfaction performance_score recruitment_source department employment_status term_reason dateof_hire race_desc hispanic_latino -Band marital_desc sex dob zip state pay_rate from_diversity_job_fair_id perf_score_id dept_id emp_status_id gender_id marital_status_id married_id emp_id employee_name manager_name manager_id dateof_termination -100 Missing Rows Univariate Distribution Histogram emp_satisfaction marital_status_id engagement_survey perf_score_id number_of_subgroups position_id 100-50 special_projects_count 100-100-Bar Chart (with frequency) ## 3 columns ignored with more than 50 categories. ## employee_name: 196 categories ## dob: 195 categories ## dateof_hire: 79 categories manager_name Ketsia Liebig David Stanley Kelley Spirea Brannon Miller Janet King Elijiah Gray Peter Monroe Lynn Daneault Michael Albert John Smith Kissy Sullivan Simon Roup Brian Champaigne Amy Dunn Webster Butler Jennifer Zamora Alex Sweetwater Brandon R. LeBlanc Eric Dougall Debra Houlihan Board of Directors -MA-CT-VTX-WA-UTI-OR-NY-NH-ND-NT-ME-KIN-ID-GA-ALmarital_desc citizen_desc US Citizen -Married -Eligible NonCitizen -Non-Citizen -Widowed hispanic_latino race_desc term_reason Black or African American -N/A - still employed -Two or more races yes -American Indian or Alaska Native Hispanic -Frequency Page 1 employment_status recruitment_source department Employee Referral Professional Society Pay Per Click - Google Monster.com MBTA ads Diversity Job Fair Website Banner Ads Newspager/Magazine Billboard Search Engine - Google Bing Yahoo On-campus Recruiting Vendor Referral Indeed Glassdoor Word of Mouth Other Internet Search Social Networks - Facebook Twitter etc Information Session Careerbuilder -Production Active -Software Engineering -Leave of Absence -Admin Offices -Executive Office -100 performance_score last_performance_review_date married_id Fully Meets -Exceeds -PIP -Needs Improvement from_diversity_job_fair_id emp_status_id gender_id Frequency Page 2



Correlation Analysis

Principal Component Analysis

Insufficient complete rows! Skipping correlation analysis.

Insufficient complete rows! Skipping principal component analysis.