Project 2

# Excel File

# Data Manipulation

## Create a new variable that reflects monthly income and add it to the data set as a new column

Created in the attached spreadsheet in column MonthlyIncome.

## Compute age for each employee based on the DOB variable. Add this to the data set as a new column

Created in the attached spreadsheet in column Age.

## Apply the codes to the PerformanceScore variable and verify that you get the same values as those that appear in PerfScoreID variable

Analysis completed, there was one employee, Alex Forrest, who had a PIP but was assigned a PerfScoreID of 3.

# Analyses

## Report the descriptive statistics along with the frequency distribution and provide a detailed interpretation of how you would characterize the salary variable.

A screenshot of a paper

Description automatically generated

R Studio descriptive statistics:

> library(tidyverse)

> #import data set

> library(readxl)

> Project2\_JW <- read\_excel("Library/CloudStorage/Box-Box/2024 - Fall/9050 - Research Methods/Projects/Project 2/Project2\_JW.xlsx")

> View(Project2\_JW)

>

> #summary statistics

> summary(Project2\_JW$Salary)

Min. 1st Qu. Median Mean 3rd Qu. Max.

45046 55502 62810 69021 72036 250000

>

> #mean, median, mode

> mean(Project2\_JW$Salary)

[1] 69020.68

>

> #median

> median(Project2\_JW$Salary)

[1] 62810

>

> #mode

> mode = function(){

+ return(sort(-table(Project2\_JW$Salary))[1])

+ }

> mode()

57815

-2

>

> #frequency table

> table <- table(Project2\_JW$Salary)

> print(table)

45046 45069 45115 45395 45433 45998 46120 46335 46428 46430 46654 46664

1 1 1 1 1 1 1 1 1 1 1 1

46738 46799 46837 46998 47001 47211 47414 47434 47750 47837 47961 48285

1 1 1 1 1 1 1 1 1 1 1 1

48413 48495 48513 48888 49256 49773 49920 50178 50274 50373 50428 50470

1 1 1 1 1 1 1 1 1 1 1 1

50482 50750 50825 50923 51044 51259 51337 51505 51777 51908 51920 52057

1 1 1 1 1 1 1 1 1 1 1 1

52087 52177 52249 52505 52599 52624 52674 52788 52846 52984 53018 53060

1 1 1 1 1 1 1 1 1 1 1 1

53171 53180 53189 53250 53366 53492 53564 54005 54132 54237 54285 54670

1 1 1 1 1 1 1 1 1 1 1 1

54828 54933 55000 55140 55315 55425 55578 55688 55722 55800 55875 55965

1 1 1 1 1 1 1 1 1 1 1 1

56147 56149 56294 56339 56847 56991 57568 57575 57583 57748 57815 57834

1 1 1 1 1 1 1 1 1 1 2 1

57859 57954 57975 58062 58207 58273 58275 58370 58371 58523 58530 58709

1 1 1 1 1 1 1 1 1 1 1 1

58939 59026 59124 59144 59231 59238 59365 59369 59370 59472 59728 59892

1 1 1 1 1 1 1 1 1 1 1 1

60070 60120 60270 60340 60380 60436 60446 60627 60656 60724 60754 61154

1 1 1 1 1 1 1 1 1 1 1 1

61242 61349 61355 61422 61555 61568 61584 61656 61729 61809 61844 61962

2 1 1 1 1 1 1 1 1 1 1 1

62061 62065 62068 62162 62385 62425 62506 62514 62659 62810 62910 62957

1 1 1 1 1 1 1 1 1 1 1 1

63000 63003 63025 63051 63108 63291 63322 63353 63381 63430 63450 63478

1 1 2 1 1 1 1 1 1 1 1 1

63515 63676 63682 63695 63763 63813 63878 63973 64021 64057 64066 64246

1 1 1 1 1 1 1 1 1 1 1 1

64375 64397 64520 64724 64738 64786 64816 64919 64955 64971 64991 64995

1 1 1 1 1 1 1 1 1 1 1 1

65288 65310 65707 65714 65729 65893 65902 66074 66149 66441 66541 66593

1 1 1 1 1 1 1 1 1 1 1 1

66738 66808 66825 67176 67237 67251 68051 68099 68182 68407 68678 68829

1 1 1 1 1 1 1 1 1 1 1 1

68898 68999 69340 70131 70187 70468 70507 70545 70621 71339 71707 71776

1 1 1 1 1 1 1 1 1 1 1 1

71860 71966 72106 72202 72460 72609 72640 72992 73330 74226 74241 74312

1 1 1 1 1 1 1 1 1 1 1 1

74326 74417 74669 74679 74813 75188 75281 76029 77692 77915 80512 81584

1 1 1 1 1 1 1 1 1 1 1 1

82758 83082 83363 83552 83667 84903 85028 86214 87565 87826 87921 88527

1 1 1 1 1 1 1 1 1 1 1 1

88976 89292 89883 90100 92328 92329 92989 93046 93093 93206 93396 93554

1 1 1 1 1 1 1 1 1 1 1 1

95660 95920 96820 97999 99020 99280 99351 100031 100416 101199 103613 104437

1 1 1 1 1 1 1 1 1 1 1 1

105688 105700 106367 107226 108987 110000 110929 113999 114800 120000 138888 140920

1 1 1 1 1 1 1 1 1 1 1 1

148999 150290 157000 170500 178000 180000 220450 250000

1 1 1 1 1 1 1 1

> barplot(table)

A black line with numbers

Description automatically generated with medium confidence

Plot:

A graph with red dots

Description automatically generated

# Standardize the Salary Variable

## Which employee (ID number) has the largest z-score on Salary and what is the z-score for this person? Which employee (ID number) has the smallest z-score on Salary and what is the z-score for this person?

Janet King has the largest z-score, which is 7.1940981. Employee ID 10271 has the lowest z-score, which is -0.9530163.

## Compute descriptive statistics for the standardized Salary variable. Report your results and produce a frequency distribution for the standardized Salary scores. Compare this distribution to the one you produced in Question 1. Are they the same or different? Explain using both your graphical results and words.

Summary Statistics:

> #summary statistics

> summary(Project2\_JW$SalaryZscore)

Min. 1st Qu. Median Mean 3rd Qu. Max.

-0.9530 -0.5374 -0.2469 0.0000 0.1199 7.1941

> #mean

> mean(Project2\_JW$SalaryZscore)

[1] 1.285146e-17

>

> #median

> median(Project2\_JW$SalaryZscore)

[1] -0.2468806

> #mode

> mode = function(){

+ return(sort(-table(Project2\_JW$SalaryZscore))[1])

+ }

> mode()

-0.445436523125001

-2

### Frequency Table:

> #frequency table

> table <- table(Project2\_JW$SalaryZscore)

> print(table)

-0.953016293652753 -0.952102022000929 -0.950273478697282

1 1 1

-0.939143215109861 -0.937632679337283 -0.915173397455524

1 1 1

-0.910323782606719 -0.901777330209236 -0.898080492660557

1 1 1

-0.89800099077779 -0.889096779907853 -0.888699270494017

1 1 1

-0.885757700831627 -0.883332893407225 -0.881822357634646

1 1 1

-0.87542245607188 -0.875303203247729 -0.866955505557164

1 1 1

-0.858886064456284 -0.858091045628611 -0.84552974815138

1 1 1

-0.842071416251002 -0.83714229951943 -0.82426299451113

1 1 1

-0.819174874014023 -0.815915296820565 -0.815199779875659

1 1 1

-0.800293176856792 -0.785664830427611 -0.765113593732268

1 1 1

-0.759270205348872 -0.749014462471892 -0.745198372099062

1 1 1

-0.741263028902081 -0.73907672712598 -0.737407187587867

1 1 1

-0.736930176291264 -0.726276924000447 -0.723295603396674

1 1 1

-0.719400011141077 -0.714590147233656 -0.706043694836172

1 1 1

-0.702943121408248 -0.696264963255796 -0.685452707199445

1 1 1

-0.680245333878187 -0.679768322581583 -0.674322443612024

1 1 1

-0.673129915370515 -0.669552330645987 -0.666690262866365

1 1 1

-0.656514021872152 -0.652777433382089 -0.651783659847498

1 1 1

-0.649796112778316 -0.64526450546058 -0.642958950860329

1 1 1

-0.637473320949386 -0.636121788942342 -0.634452249404229

1 1 1

-0.630039894910645 -0.629682136438192 -0.629324377965739

1 1 1

-0.626899570541337 -0.622288461340834 -0.617279842726495

1 1 1

-0.614417774946873 -0.596887609796686 -0.591839240240963

1 1 1

-0.58766539139568 -0.585757346209265 -0.570453233776562

1 1 1

-0.564172585037946 -0.559998736192664 -0.55733542311996

1 1 1

-0.551770291326249 -0.544813876584112 -0.540441273031911

1 1 1

-0.534359379000213 -0.529986775448012 -0.528635243440969

1 1 1

-0.525534670013044 -0.522553349409271 -0.518975764684743

1 1 1

-0.51174109335292 -0.511661591470153 -0.505897704969524

1 1 1

-0.50410891260726 -0.483915434384369 -0.478191298825124

1 1 1

-0.455255005646762 -0.454976749057076 -0.454658741526007

1 1 1

-0.448099836197706 -0.445436523125001 -0.444681255238712

1 2 1

-0.443687481704121 -0.439911142272675 -0.439076372503618

1 1 1

-0.435618040603241 -0.429854154102613 -0.427230591971292

1 1 1

-0.427151090088525 -0.423374750657079 -0.423334999715695

1 1 1

-0.417292856625381 -0.417014600035696 -0.409899181528024

1 1 1

-0.400756465009786 -0.397298133109408 -0.393402540853811

1 1 1

-0.392607522026138 -0.389149190125761 -0.388870933536076

1 1 1

-0.383822563980353 -0.383663560214819 -0.383623809273435

1 1 1

-0.379569213252303 -0.36939297225809 -0.362873817871173

1 1 1

-0.355798150304884 -0.353810603235702 -0.347847962028155

1 1 1

-0.3450653961313 -0.343475358475954 -0.34124930575847

1 1 1

-0.340851796344634 -0.333656875954194 -0.332504098654069

1 1 1

-0.329801034639981 -0.328608506398471 -0.312708129845014

1 1 1

-0.309210047003253 -0.304956696275203 -0.304718190626901

2 1 1

-0.302054877554197 -0.296768002350173 -0.296251240112185

1 1 1

-0.295615225050047 -0.292753157270424 -0.289851338549418

1 1 1

-0.286671263238727 -0.285279980290299 -0.280589369207029

1 1 1

-0.276654026010049 -0.276495022244514 -0.276375769420363

1 1 1

-0.2726391809303 -0.263774721001748 -0.262184683346402

1 1 1

-0.258964857094327 -0.258646849563258 -0.252882963062629

1 1 1

-0.246880570913699 -0.242905476775335 -0.241037182530303

1 1 1

-0.239327892050807 -0.239208639226656 -0.238334118516216

1 1 2

-0.237300594040241 -0.235034790381373 -0.227760368108166

1 1 1

-0.226528088925273 -0.22529580974238 -0.224182783383638

1 1 1

-0.22223498725584 -0.221439968428167 -0.220326942069425

1 1 1

-0.21885615723823 -0.212456255675463 -0.212217750027161

1 1 1

-0.211700987789174 -0.208997923775086 -0.207010376705904

1 1 1

-0.204426565515967 -0.200650226084521 -0.198742180898106

1 1 1

-0.197311147008295 -0.196953388535842 -0.189798219086786

1 1 1

-0.184670347648296 -0.183795826937856 -0.178906461147668

1 1 1

-0.170797269105404 -0.170240755926033 -0.168332710739618

1 1 1

-0.167140182498109 -0.163045835535594 -0.161614801645783

1 1 1

-0.160978786583644 -0.160183767755971 -0.160024763990437

1 1 1

-0.148377738165029 -0.147503217454589 -0.131722093725282

1 1 1

-0.131443837135597 -0.130847573014842 -0.124328418627924

1 1 1

-0.123970660155472 -0.117133498237485 -0.114152177633711

1 1 1

-0.102544902749687 -0.098569808611323 -0.0965027596593735

1 1 1

-0.090738873158745 -0.08795630726189 -0.087280541258368

1 1 1

-0.0733279608327089 -0.0709031534083066 -0.0703466402289356

1 1 1

-0.0385458871220203 -0.0366378419356054 -0.033338513800763

1 1 1

-0.024394551989443 -0.0136220468744755 -0.00761965472554521

1 1 1

-0.00487683977007377 -0.000861994690325712 0.0126930763214969

1 1 1

0.0441360709559594 0.0463621236734435 0.0575321382022475

1 1 1

0.0590824249162096 0.0605929606887881 0.063614032233945

1 1 1

0.0921552081474015 0.106783554576583 0.109526369532054

1 1 1

0.11286544860828 0.117079048394946 0.122644180188657

1 1 1

0.126460270561486 0.136716013438467 0.14263890370463

1 1 1

0.143871182887522 0.157863514254565 0.171299332442237

1 1 1

0.206916175921982 0.207512440042737 0.210334756880975

1 1 1

0.210891270060346 0.214508605726258 0.224525842954936

1 1 1

0.224923352368773 0.230249978514181 0.245156581533048

1 1 1

0.248853419081727 0.278587123236692 0.344692938757693

1 1 1

0.353557398686245 0.456790593459569 0.499403602622835

1 1 1

0.546071207807234 0.558950512815534 0.570120527344338

1 1 1

0.577633455265847 0.582204813524966 0.63133697707515

1 1 1

0.636305844748106 0.683450461229108 0.737153983038411

1 1 1

0.747528978739542 0.751305318170988 0.775394388649477

1 1 1

0.793242561330733 0.805803858807964 0.829296665165698

1 1 1

0.837922619445949 0.926487716848708 0.926527467790091

1 1 1

0.952763089103297 0.955028892762164 0.956897187007196

1 1 1

0.961389043383547 0.96894172224644 0.975222370985055

1 1 1

1.05893785353901 1.06927309829876 1.10504894554404

1 1 1

1.15191530543535 1.19250101658805 1.2028362613478

1 1 1

1.20565857818604 1.23268921832692 1.24799333075962

1 1 1

1.27911831786301 1.37507709036313 1.40783186606325

1 1 1

1.45756029373419 1.4580373050308 1.48455118293369

1 1 1

1.51869724158224 1.58869864935884 1.62896635298047

1 1 1

1.66589497752587 1.78793036757366 1.81977087162196

1 1 1

2.02647576681691 2.77729154767118 2.85806546056274

1 1 1

3.1792133160012 3.23053178132749 3.49726059801174

1 1 1

4.03389830669094 4.33203036706827 4.41153224983555

1 1 1

6.01945782880396 7.19409814669064

1 1

Frequency Table Bar Plot:

A black line with numbers

Description automatically generated with medium confidence

Frequency Table Plot:

A graph with red dots

Description automatically generated

\*\*ANALYSIS REQUIRED\*\*

## Compute descriptive statistics for the monthly salary variable you created at the beginning of this project. Report your results and produce a frequency distribution for these scores. Compare this distribution to the ones you produced in Questions 1 & 3. Are they the same or different? Explain using both your graphical results and words.

5. Take the square root of the Salary variable. Compute descriptive statistics for this

new variable Report your results and produce a frequency distribution for the

square root scores. Compare this distribution to the ones you produced in

Questions 1, 3, & 4. Are they the same or different? Explain using both your

graphical results and words.