# Capstone Results

## Part 1: Data Collection

#### Data Processing and Modifications Log 1

The analyzed data has information on Respondent, MainBranch, Hobbyist, OpenSourcer, OpenSource, Employment, Country, Student, EdLevel, UndergradMajor, EduOther, OrgSize, DevType, YearsCode, Age1stCode, YearsCodePro, CareerSat, JobSat, MgrIdiot, MgrMoney, MgrWant, JobSeek, LastHireDate, LastInt, FizzBuzz, JobFactors, ResumeUpdate, CurrencySymbol, CurrencyDesc, CompTotal, CompFreq, ConvertedComp, WorkWeekHrs, WorkPlan, WorkChallenge, WorkRemote, WorkLoc, ImpSyn, CodeRev, CodeRevHrs, UnitTests, PurchaseHow, PurchaseWhat, LanguageWorkedWith, LanguageDesireNextYear, DatabaseWorkedWith, DatabaseDesireNextYear, PlatformWorkedWith, PlatformDesireNextYear, WebFrameWorkedWith, WebFrameDesireNextYear, MiscTechWorkedWith, MiscTechDesireNextYear, DevEnviron, OpSys, Containers, BlockchainOrg, BlockchainIs, BetterLife, ITperson, OffOn, SocialMedia, Extraversion, ScreenName, SOVisit1st, SOVisitFreq, SOVisitTo, SOFindAnswer, SOTimeSaved, SOHowMuchTime, SOAccount, SOPartFreq, SOJobs, EntTeams, SOComm, WelcomeChange, SONewContent, Age, Gender, Trans, Sexuality, Ethnicity, Dependents, SurveyLength, SurveyEase.

#### Data Processing and Modifications Log 2

The DataFrame has 11552 rows at the start.

Data Processing and Modifications Log 3

There are 154 duplicates

#### Data Processing and Modifications Log 4

All Duplicates were deleted.

#### Data Processing and Modifications Log 5

The DataFrame has 11398 rows after deleting duplicates.

#### Data Processing and Modifications Log 6

repaired workloc column

#### Data Processing and Modifications Log 7

There are 30967 missing values in columns which can´t be replaced

#### Data Processing and Modifications Log 8

ALL rows with missing values deleted

#### Data Processing and Modifications Log 9

The DataFrame has 1008 rows after deleting missing rows.

#### Data Processing and Modifications Log 10

There are 3 different compensation categories. These include ['Monthly' 'Yearly' 'Weekly']

#### Data Processing and Modifications Log 11

A normalized Annual Compensation column was created to reflect a comparable income of responders

## Part 2 Data wrangling and exploratory analysis

#### Data Analysis Log 1

The median age is 29.0 years old!

#### Data Analysis Log 2

The median annual salary is 57745.0 dollars!

#### Data Analysis Log 3

There were 10480 Men!

#### Data Analysis Log 4

The female median annual salary is 57708.0 dollars!

#### Data Analysis Log 5

Age summary

count 11111.000000  
mean 30.778895  
std 7.393686  
min 16.000000  
25% 25.000000  
50% 29.000000  
75% 35.000000  
max 99.000000

#### Data Analysis Log 6

Salary summary

count 1.058200e+04  
mean 1.315967e+05  
std 2.947865e+05  
min 0.000000e+00  
25% 2.686800e+04  
50% 5.774500e+04  
75% 1.000000e+05  
max 2.000000e+06

#### Data Analysis Log 7

Half of Responders make between 100000.0 and 26868.0 dollars yearly.

#### Data Analysis Log 8

Special cases(outliers) are People who make over 209698.0 and under -82830.0 dollars yearly

#### Data Analysis Log 9

Salaries data considered will disregard outliers, the data analyzed is summarized below:

count 9703.000000  
mean 59883.208389  
std 43394.336755  
min 0.000000  
25% 24060.000000  
50% 52704.000000  
75% 85574.500000  
max 209356.000000  
Name: ConvertedComp, dtype: float64

#### Data Analysis Log 10

The median annual salary of the analyzed data is 52704.0 dollars!

#### Data Analysis Log 11

The analyzed data shows that the average annual salary is 59883.20838915799 dollars!

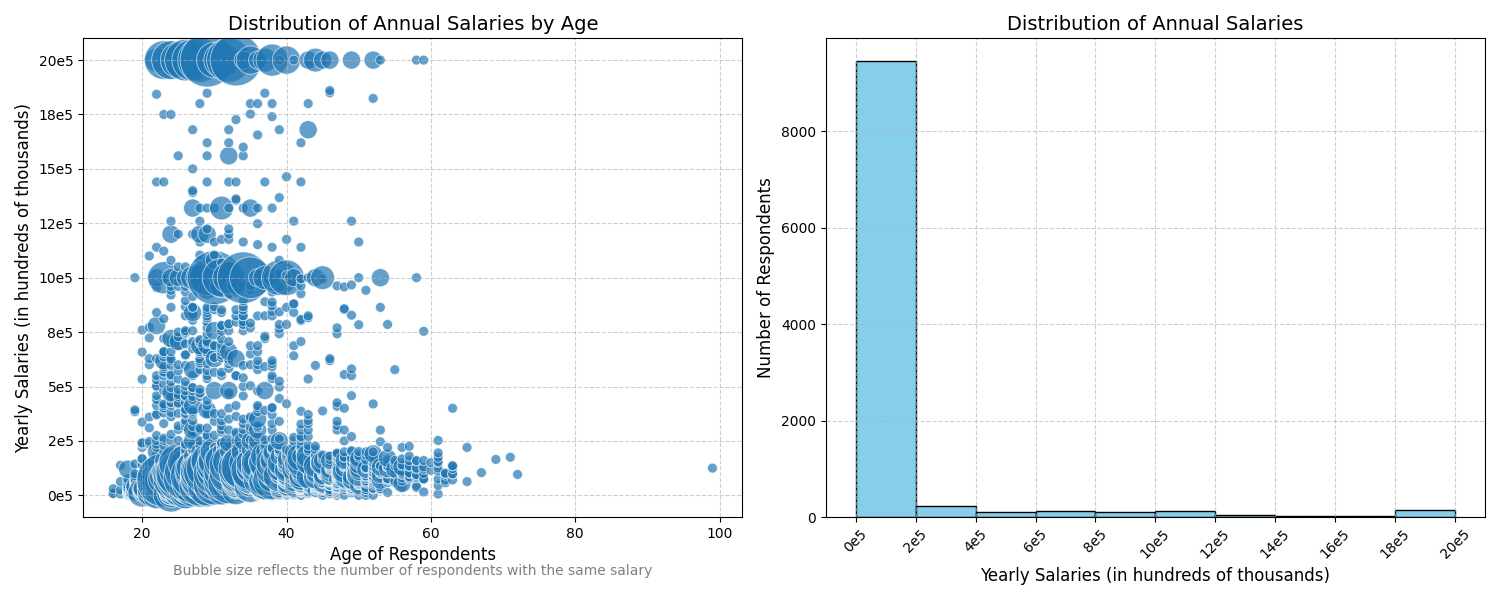
#### Data Analysis Log 12

Age-determined/correlated factors:  
 Correlation  
ConvertedComp 0.401821  
Age 1.000000

CodeRevHrs -0.012878  
Respondent 0.002180  
CompTotal 0.006337  
WorkWeekHrs 0.032032

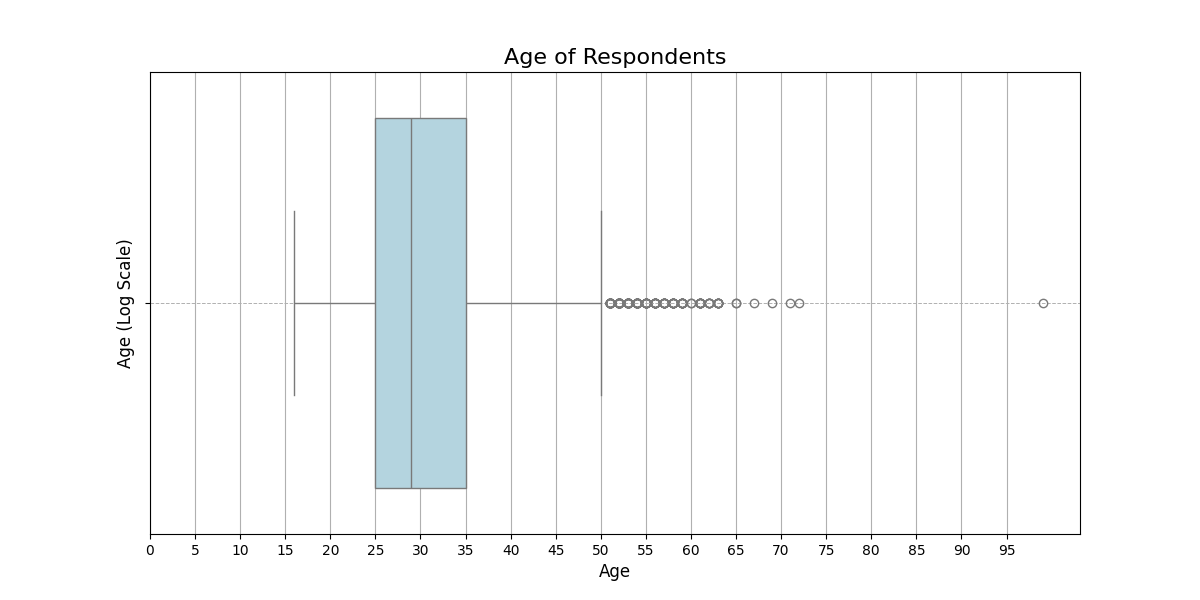
### Graph 2

Description of Graph 2



### Graph 3

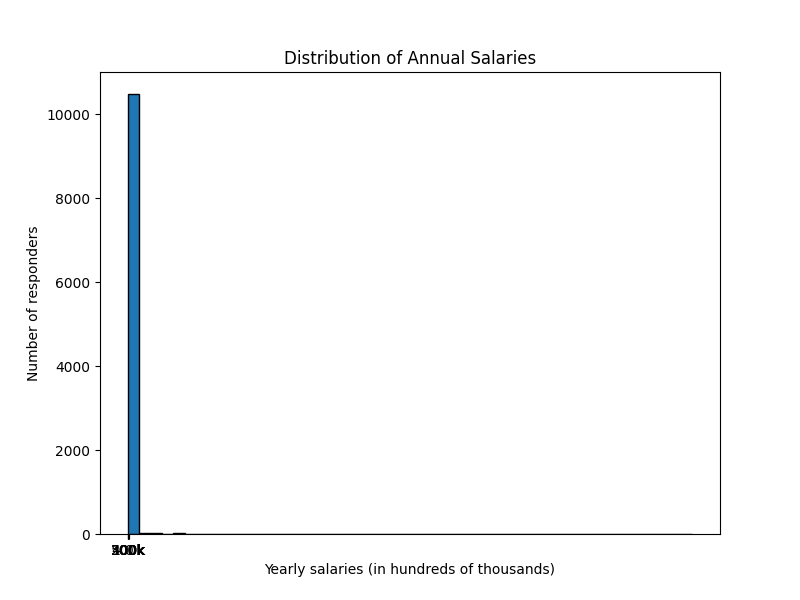
Description of Graph 3



## Part 3 Data visualization

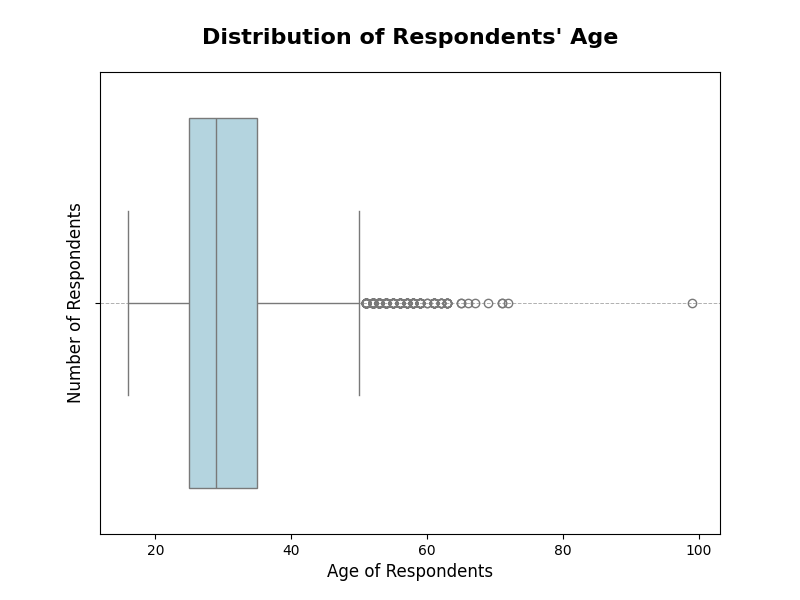
#### Graph 1

Description of Graph 1



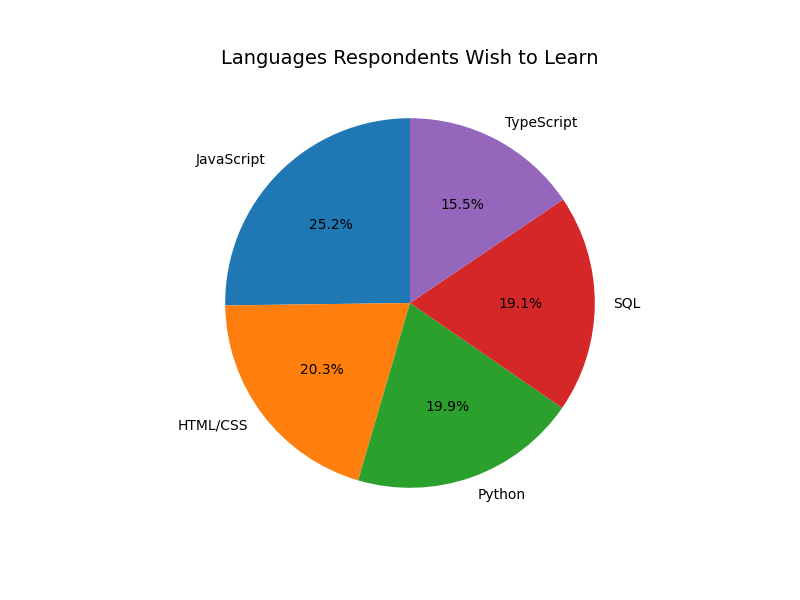
#### Graph 2

Description of Graph 2



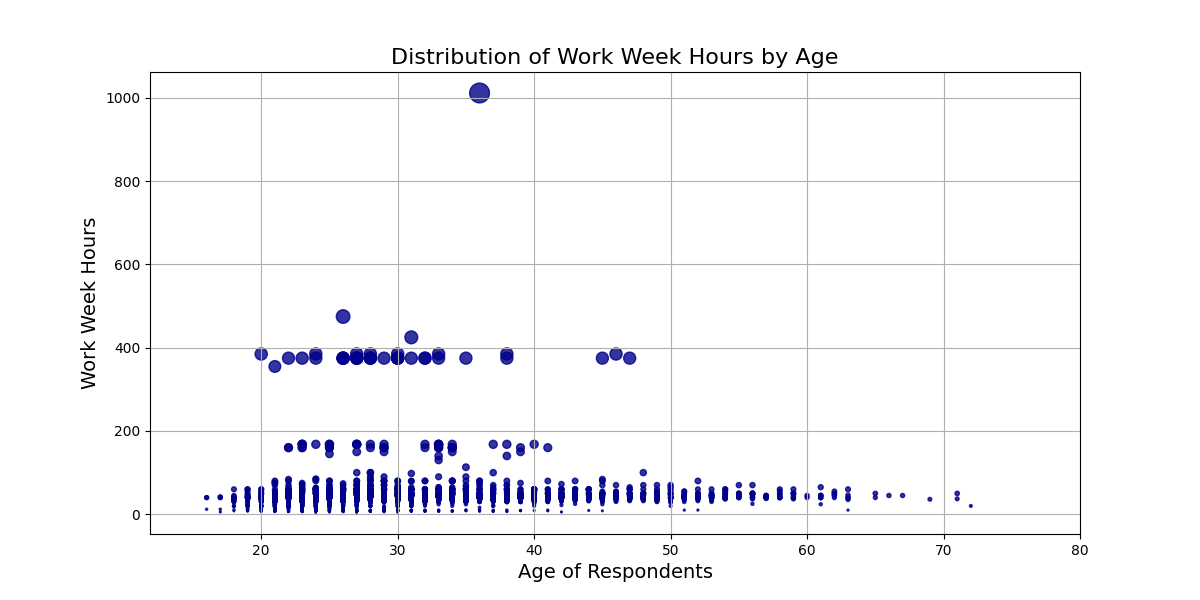
#### Graph 3

Description of Graph 3



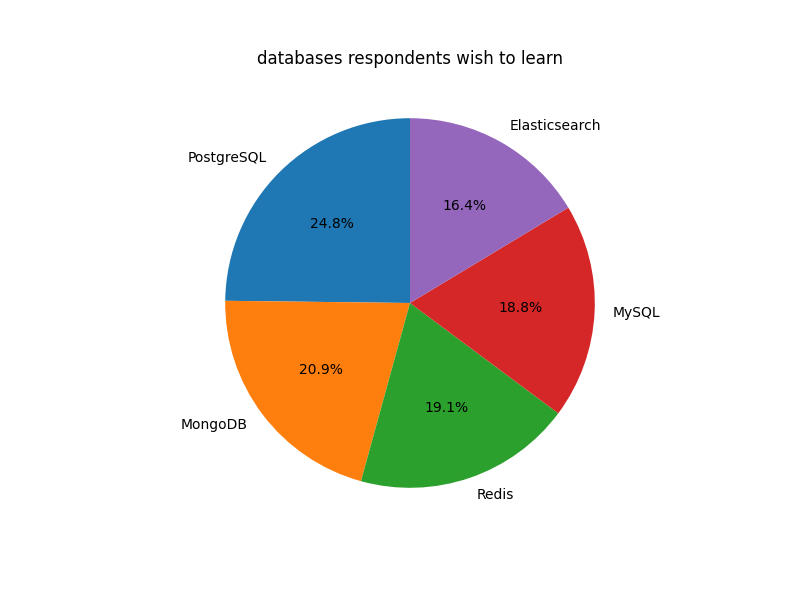
#### Graph 4

Description of Graph 4



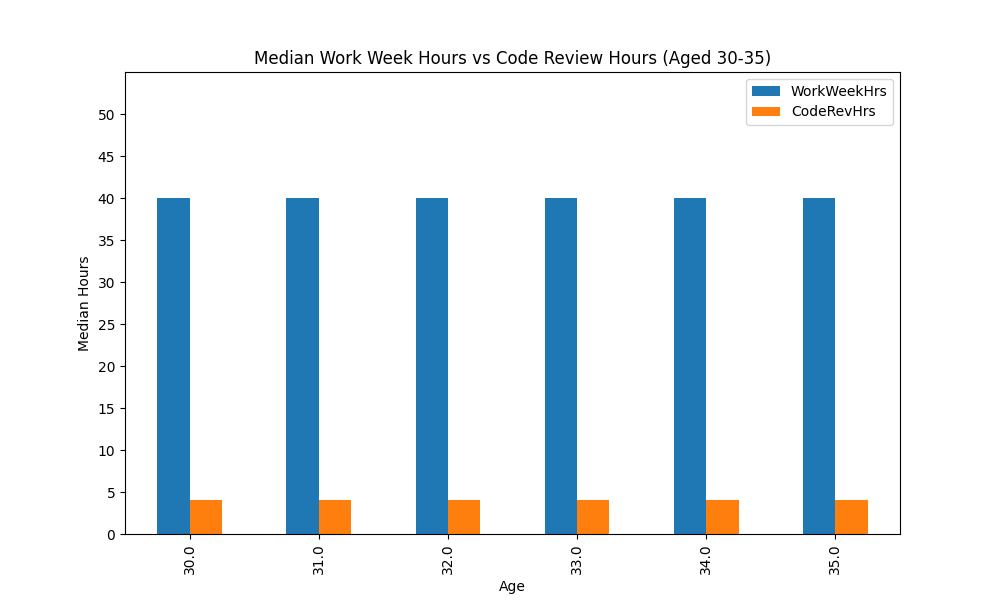
#### Graph 5

Description of Graph 5



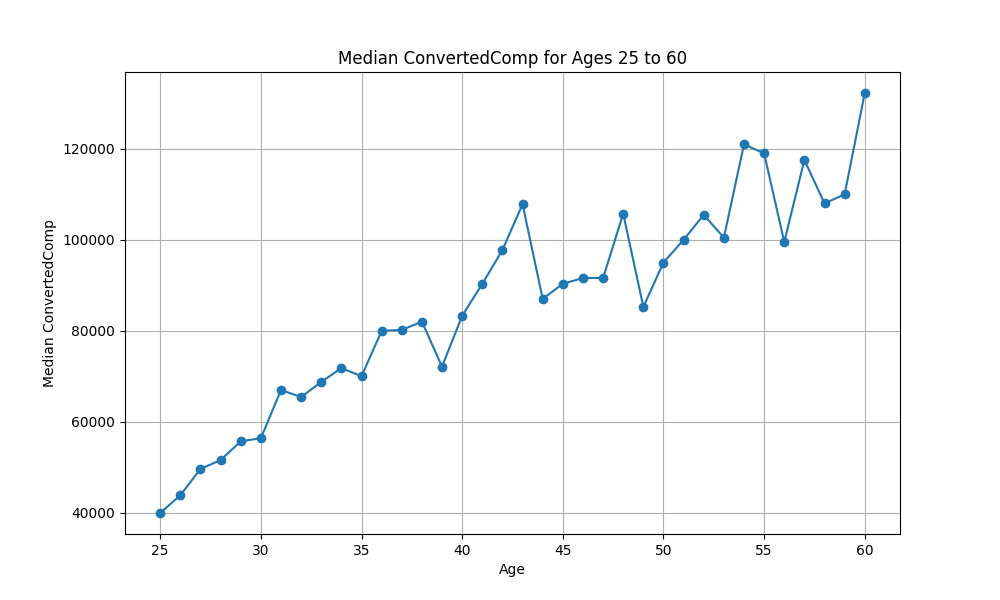
#### Graph 6

Description of Graph 6



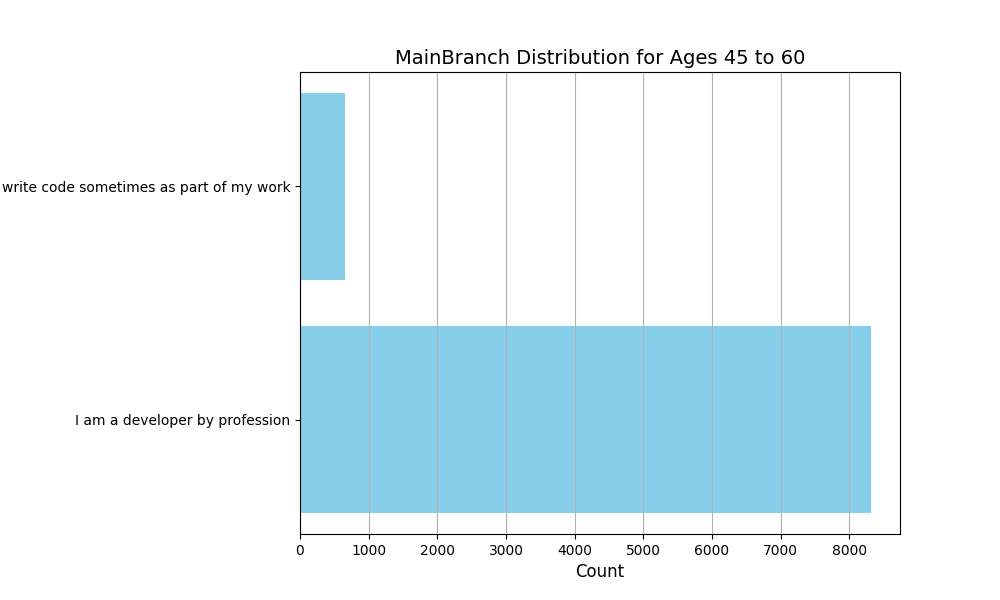
#### Graph 7

Description of Graph 7



#### Graph 8

Description of Graph 8



#### Graph 9

Description of Graph 9

Ein Bild, das Text, Screenshot, Diagramm, Reihe enthält.

Automatisch generierte Beschreibung

## Part 4 Dashboard results

View original under: <https://lookerstudio.google.com/s/uuW9w-kMkRk>

#### Ein Bild, das Text, Screenshot, Software, Display enthält. Automatisch generierte Beschreibung

Ein Bild, das Text, Screenshot, Software, Diagramm enthält.

Automatisch generierte BeschreibungEin Bild, das Text, Screenshot, Multimedia-Software, Software enthält.

Automatisch generierte Beschreibung