

# StackOverflow Annual Developer Survey 2020 Analysis

Fechar Ourotcha

11/10/2020

## Overview:

Every year, StackOverflow deploys a survey to many of its participants. With nearly 65,000 responses fielded from over 180 countries and dependent territories, StackOverflow's 2020 Annual Developer Survey examines all aspects of the developer experience from career satisfaction and job search to education and opinions on open-source software.

In this analysis, we are going to analyze the survey dataset and draw key insights about the participant in regards to their race, gender, age, compensation, etc. The purpose of this analysis is to provide a detailed insight that will be useful for aspiring developers who have not begun their journey yet.

The Original dataset contains sixty-one variables and 64461 observations, however, we will be using only 11 of those variables.

```
# Loading the necessary packages and Importing the data
```

```
library(tidyverse)
```

```
## Warning: package 'tidyverse' was built under R version 3.6.3
```

```
## -- Attaching packages ----- tidyverse 1.3.0 --
```

```
## v ggplot2 3.2.1    v purrr   0.3.3
## v tibble  2.1.3    v dplyr   0.8.3
## v tidyr   1.0.2    v stringr 1.4.0
## v readr   1.3.1    v forcats 0.4.0
```

```
## -- Conflicts ----- tidyverse_conflicts() --
## x dplyr::filter() masks stats::filter()
## x dplyr::lag()     masks stats::lag()
```

```
library(readxl)
```

```
survey_2020 <- read.csv("Survey_Results_Dataset.csv")
```

```
# Examining the structure of the dataset and filtering for variables that we are using
```

```
survey_2020_filtered <- survey_2020 %>%
  select(WorkWeekHrs, YearsCode, Age, Gender, Ethnicity, EdLevel,
         Age1stCode)
```

```
glimpse(survey_2020_filtered)
```

```
## Observations: 64,461
## Variables: 11
## $ WorkWeekHrs    <dbl> 50, NA, NA, 40, NA, NA, NA, 39, 50, 36, 40, 40, 36, 4...
## $ YearsCode      <fct> 36, 7, 4, 7, 15, 6, 6, 17, 6, 8, 10, 7, 35, 5, 4, 37,...
## $ Age            <dbl> NA, NA, NA, 25, 31, NA, NA, 36, 30, 22, 23, 49, 53, 2...
## $ Gender         <fct> Man, NA, NA, Man, Man, Man, NA, Man, Man, Man, Man, M...
## $ Ethnicity      <fct> White or of European descent, NA, NA, White or of Eur...
## $ EdLevel        <fct> "Masterâ\200\231s degree (M.A., M.S., M.Eng., MBA, et...
## $ DevType        <fct> "Developer, desktop or enterprise applications;Develo...
## $ Country        <fct> Germany, United Kingdom, Russian Federation, Albania,...
## $ ConvertedComp  <dbl> NA, NA, NA, NA, NA, NA, NA, NA, 116000, NA, 32315, 40070,...
## $ MainBranch     <fct> "I am a developer by profession", "I am a developer b...
## $ Age1stCode     <fct> 13, 19, 15, 18, 16, 14, 18, 12, 20, 14, 13, 42, 14, 1...
```

```
head(survey_2020_filtered)
```

```
##   WorkWeekHrs YearsCode Age Gender Ethnicity
## 1          50        36  NA   Man White or of European descent
## 2          NA         7  NA   <NA> <NA>
## 3          NA         4  NA   <NA> <NA>
## 4          40         7  25   Man White or of European descent
## 5          NA        15  31   Man White or of European descent
## 6          NA         6  NA   Man White or of European descent
##                                     EdLevel
## 1                                     Masterâ\200\231s degree (M.A., M.S., M.Eng., MBA, etc.)
## 2                                     Bachelorâ\200\231s degree (B.A., B.S., B.Eng., etc.)
## 3                                     <NA>
## 4                                     Masterâ\200\231s degree (M.A., M.S., M.Eng., MBA, etc.)
## 5                                     Bachelorâ\200\231s degree (B.A., B.S., B.Eng., etc.)
## 6 Secondary school (e.g. American high school, German Realschule or Gymnasium, etc.)
##                                     DevType
## 1 Developer, desktop or enterprise applications;Developer, full-stack
## 2 Developer, full-stack;Developer, mobile
## 3 <NA>
## 4 <NA>
## 5 <NA>
## 6 Designer;Developer, front-end;Developer, mobile
##   Country ConvertedComp
## 1   Germany          NA
## 2 United Kingdom          NA
## 3 Russian Federation          NA
## 4   Albania          NA
## 5 United States          NA
## 6   Germany          NA
##                                     MainBranch Age1stCode
## 1                                     I am a developer by profession          13
## 2                                     I am a developer by profession          19
## 3                                     I code primarily as a hobby            15
## 4                                     I am a developer by profession          18
## 5 I used to be a developer by profession, but no longer am            16
## 6                                     I am a developer by profession          14
```

```
tail(survey_2020_filtered)
```

```
##      WorkWeekHrs YearsCode Age Gender Ethnicity
## 64456          NA      <NA>  NA  <NA>      <NA>
## 64457          NA        10  NA  <NA>      <NA>
## 64458          NA      <NA>  NA  <NA>      <NA>
## 64459          NA      <NA>  NA  <NA>      <NA>
## 64460          NA      <NA>  NA  <NA>      <NA>
## 64461          NA      <NA>  NA  <NA>      <NA>
##                                     EdLevel      DevType
## 64456                                     <NA>      <NA>
## 64457 Masterâ\200\231s degree (M.A., M.S., M.Eng., MBA, etc.) Senior executive/VP
## 64458                                     <NA>      <NA>
## 64459                                     Primary/elementary school      <NA>
## 64460                                     <NA>      <NA>
## 64461      Other doctoral degree (Ph.D., Ed.D., etc.)      <NA>
##      Country ConvertedComp MainBranch Age1stCode
## 64456  Netherlands          NA      <NA>      <NA>
## 64457  United States          NA      <NA>        16
## 64458    Morocco          NA      <NA>      <NA>
## 64459   Viet Nam          NA      <NA>      <NA>
## 64460    Poland          NA      <NA>      <NA>
## 64461    Spain          NA      <NA>      <NA>
```

## Data Cleaning & EDA: