

# 2021 New Coder Survey

This project will take a deep delve into a dataset containing the answers to 49 questions asked to new software developers and learners. We will analyse the results and extract meaningful information to hopefully gain insight into how developers are feeling about the Data Science field, as well as the habits and expectations of new and would-be programmers.

This will be achieved using built-in Python libraries, namely Numpy, Pandas, and Seaborn which can give us the tools we need to manipulate and visualise our data in easy to understand and meaningful ways.

This project is the final part of the Jovian course: [Data Analysis with Python: Zero to Pandas](#).

## Downloading the Dataset

To begin with, we will use the Python Library 'opendatasets' to download the dataset from Kaggle. When we do this, we will be asked for a Kaggle username and key which can be obtained from the Kaggle website.

```
!pip install jovian opendatasets --upgrade --quiet
```

Let's begin by downloading the data, and listing the files within the dataset.

```
dataset_url = 'https://www.kaggle.com/datasets/fccuser/2021-new-coder-survey'
```

```
import opendatasets as od
od.download(dataset_url)
```

Please provide your Kaggle credentials to download this dataset. Learn more:

<http://bit.ly/kaggle-creds>

Your Kaggle username: adammason2905

Your Kaggle Key: .....

Downloading 2021-new-coder-survey.zip to ./2021-new-coder-survey

100%|██████████| 1.82M/1.82M [00:00<00:00, 13.8MB/s]

The dataset has been downloaded and extracted.

```
# Change this
data_dir = './2021-new-coder-survey'
```

```
import os
os.listdir(data_dir)
```

```
['2021 New Coder Survey.csv']
```

Let us save and upload our work to Jovian before continuing.

```
project_name = "2021-new-coder-survey" # change this (use lowercase letters and hyphens)
```

```
!pip install jovian --upgrade -q
```

```
import jovian
```

```
jovian.commit(project=project_name)
```

```
[jovian] Updating notebook "adammasongh/2021-new-coder-survey" on https://jovian.com  
[jovian] Committed successfully! https://jovian.com/adammasongh/2021-new-coder-survey  
'https://jovian.com/adammasongh/2021-new-coder-survey'
```

## Data Preparation and Cleaning

Data preparation is the process of cleaning, transforming and restructuring our data so that we are able to use it for analysis and visualisation. When working with large datasets, this can often be a very lengthy and time-consuming process, but it is absolutely essential to perform this task to make sure we can extract useful information out of our datasets.

We will begin by importing the Pandas and Numpy libraries.

```
import pandas as pd  
import numpy as np
```

```
survey_raw_df = pd.read_csv('2021-new-coder-survey/2021 New Coder Survey.csv')
```

```
/opt/conda/lib/python3.9/site-packages/IPython/core/interactiveshell.py:3444:
```

```
DtypeWarning: Columns (8,9) have mixed types.Specify dtype option on import or set  
low_memory=False.
```

```
exec(code_obj, self.user_global_ns, self.user_ns)
```

```
survey_raw_df
```

Timestamp	1. What is your biggest reason for learning to code?	2. What methods have you used to learn about coding? Please select all that apply.	3. Which online learning resources have you found helpful? Please select all that apply.	4. If you have attended in-person coding-related events before, which ones have you found helpful? Please select all that apply.	5. If you have listened to coding-related podcasts before, which ones have you found helpful? Please select all that apply.	6. If you have watched coding-related YouTube videos before, which channels have you found helpful? Please select all that apply.	7. About how many hours do you spend learning each week?	8. ma hav pro!
-----------	--	--	--	--	---	---	--	----------------

---

		Timestamp	1. What is your biggest reason for learning to code?	2. What methods have you used to learn about coding? Please select all that apply.	3. Which online learning resources have you found helpful? Please select all that apply.	4. If you have attended in-person coding-related events before, which ones have you found helpful? Please select all that apply.	5. If you have listened to coding-related podcasts before, which ones have you found helpful? Please select all that apply.	6. If you have watched coding-related YouTube videos before, which channels have you found helpful? Please select all that apply.	7. About how many hours do you spend learning each week?	8. ma hav pro
0		7/1/2021 10:10:23	To succeed in current career	Online resources, Books, In-person bootcamps, ...	freeCodeCamp, Mozilla Developer Network (MDN),...	conferences, workshops, Meetup.com events	The Changelog	CS Dojo, freeCodeCamp	4.0	
1		7/1/2021 10:31:01	To change careers	Online resources, Books, Online bootcamps	freeCodeCamp, Mozilla Developer Network (MDN),...	I haven't attended any in-person coding-relate...	The Changelog, Code Newbie Podcast	Adrian Twarog, Code with Ania Kubów, Coder Cod...	10.0	
2		7/1/2021 10:42:31	To change careers	Online resources, Books, Hackathons, Meetup.co...	freeCodeCamp, Mozilla Developer Network (MDN),...	Meetup.com events	I haven't listened to any podcasts	AmigosCode, Dev Ed, freeCodeCamp, Kevin Powell...	30.0	
3		7/1/2021 11:06:43	As a hobby	Online resources, Books	freeCodeCamp, Mozilla Developer Network (MDN),...	I haven't attended any in-person coding-relate...	Darknet Diaries, Real Python Podcast	freeCodeCamp, Traversy Media	NaN	
4		7/1/2021 11:14:31	To start your first career	Online resources, Books, Online bootcamps	freeCodeCamp, Stack Overflow, Coursera, Udemy	I haven't attended any in-person coding-relate...	Talk Python to Me	freeCodeCamp, The Net Ninja, Traversy Media	2.0	
...		...	...	...	...	...	...	...	...	
18121		9/29/2021 7:38:11	To change careers	Online resources, Books, In-person bootcamps	Stack Overflow, EdX, Khan Academy, Udacity, Ci...	conferences, workshops	Le code a changé (RadioFrance app)	Hashtag Programação, Curso em Video, Boson Tre...	40.0	
18122		9/29/2021 12:52:59	To change careers	Online resources, In-person bootcamps	freeCodeCamp, Mozilla Developer Network (MDN),...	I haven't attended any in-person coding-relate...	I haven't listened to any podcasts	freeCodeCamp	NaN	
18123		9/29/2021 14:18:31	To start your first career	Online resources, Books	freeCodeCamp, Mozilla Developer Network (MDN),...	I haven't attended any in-person coding-relate...	I haven't listened to any podcasts	Ben Awad, Coder Coder, Coding Train, Dev Ed, f...	10.0	

		1. What is your biggest reason for learning to code?	2. What methods have you used to learn about coding? Please select all that apply.	3. Which online learning resources have you found helpful? Please select all that apply.	4. If you have attended in-person coding-related events before, which ones have you found helpful? Please select all that apply.	5. If you have listened to coding-related podcasts before, which ones have you found helpful? Please select all that apply.	6. If you have watched coding-related YouTube videos before, which channels have you found helpful? Please select all that apply.	7. About how many hours do you spend learning each week?	8. About how many months have you been programming?
18124	9/30/2021 4:04:32	To change careers	Online resources, Books, In-person bootcamps, ...	freeCodeCamp, Stack Overflow, Codecademy, Udem...	I haven't attended any in-person coding-relate...	I haven't listened to any podcasts	AmigosCode, Code with Ania Kubów, freeCodeCamp...	6.0	
18125	10/1/2021 4:09:25	To succeed in current career	Online resources, Books	freeCodeCamp, Mozilla Developer Network (MDN),...	I haven't attended any in-person coding-relate...	I haven't listened to any podcasts	Code with Ania Kubów, Coding Train, Dev Ed, fr...	5.0	

18126 rows × 63 columns

```
print("This dataset consists of {} rows and {} columns of data.".format(survey_raw_df.s
```

This dataset consists of 18126 rows and 63 columns of data.

Before we analyse our data, we can use some built in functions to play around with our data and restructure it into the way we desire.

## Dropping columns

Firstly, we can use `.columns` to look at all of the columns in our dataframe.

```
survey_raw_df.columns
```

```
Index(['Timestamp', '1. What is your biggest reason for learning to code?',
      '2. What methods have you used to learn about coding? Please select all that apply.',
      '3. Which online learning resources have you found helpful? Please select all that apply.',
      '4. If you have attended in-person coding-related events before, which ones have you found helpful? Please select all that apply.',
      '5. If you have listened to coding-related podcasts before, which ones have you found helpful? Please select all that apply.',
      '6. If you have watched coding-related YouTube videos before, which channels have you found helpful? Please select all that apply.',
      '7. About how many hours do you spend learning each week?',
      '8. About how many months have you been programming?',
      '9. Aside from university tuition, about how much money have you spent on
```

learning to code so far (in US Dollars)?',  
'10. Are you already employed in a software development job?',  
'11. If you are already employed as a developer, is this your first software development job? ',  
'12. If you are NOT already a developer, are you interested in a software development career?',  
'13. If you are interested in a software development career, would you prefer to...',  
'14. Which of these careers are you interested in?',  
'15. When do you plan to start applying for developer jobs?',  
'16. About how much money do you expect to earn per year at your first developer job (in US Dollars)? ',  
'17. Please select up to 3 reasons why you are interested in a software development career',  
'18. After the pandemic, how many days would you ideally like to work from home versus in an office each week?',  
'19. Are you willing to relocate for a job?',  
'20. Regarding employment status, are you currently.... ',  
'21. If you are currently working, which field do you work in?',  
'22. About how much money did you earn last year from any job or employment (in US Dollars)? ',  
'23. How old are you?',  
'24. Which of the following best represents how you think of yourself?',  
'25. With which of these groups do you primarily identify?',  
'26. Which part of the world do you live in?',  
'27. If you are living in the US, which state do you currently live in? ',  
'28. About how many people live in your city?',  
'29. Is your country of citizenship different from the country where you live?',  
'30. Are you an ethnic minority in your country?',  
'31. Is English your second language?',  
'32. What is the highest degree or level of school you have completed?',  
'33. If you attended a university, what did you study?',  
'34. Are you currently a student attending or enrolled in regular school, that is in an elementary school, a middle school, a high school, a college, or a graduate school?',  
'35. What's your marital status?',  
'36. How many children do you have? By children, we mean any biological, step, or adopted children.',  
'37. Do you financially support any dependents, older relatives or relatives with disabilities?',  
'38. Do you think you have enough savings to survive for 3 months with no income?',  
'39. How much debt does your household have? [Car loans]',  
'39. How much debt does your household have? [Credit Cards]',  
'39. How much debt does your household have? [Medical Loans]',  
'39. How much debt does your household have? [Mortgages]',  
'39. How much debt does your household have? [Payday Loans]',  
'39. How much debt does your household have? [Student Loans]',  
'39. How much debt does your household have? [Other]'

```

'40. How long have you been working in your current job?',
'41. Before you got your last job, how many months did you spend looking for a
job?',
'42. If you are working, thinking about the next 12 months, how likely do you
think it is that you will lose your job or be laid off? ',
'43. If you are working, how easy would it be for you to find a job with another
employer with approximately the same income and fringe benefits you now have? ',
'44. Do you consider yourself under-employed? (Under-employment means working a
job that is below your education level.)',
'45. Please tell us how satisfied you are with each of these following aspects
of your present job [Earnings]',
'45. Please tell us how satisfied you are with each of these following aspects
of your present job [Benefits]',
'45. Please tell us how satisfied you are with each of these following aspects
of your present job [Job security]',
'45. Please tell us how satisfied you are with each of these following aspects
of your present job [Work-life balance]',
'45. Please tell us how satisfied you are with each of these following aspects
of your present job [Professional growth or leadership opportunities]',
'45. Please tell us how satisfied you are with each of these following aspects
of your present job [Workplace/company culture]',
'45. Please tell us how satisfied you are with each of these following aspects
of your present job [Diverse and inclusive work environment]',
'45. Please tell us how satisfied you are with each of these following aspects
of your present job [Weekly workload]',
'46. About how many minutes does it take you to get to work each day?',
'47. Have you served in your country's military before?',
'48. Do you currently receive disability benefits from your government?',
'49. Do you have high speed internet at your home?'],
dtype='object')

```

We have 63 columns in our dataset. To make our analysis easier, we will drop a handful of these that we won't use. Before we do this, we will create a copy of the original dataframe.

```
survey_df = survey_raw_df.copy()
```

```

columns_to_remove = ["Timestamp",
                    "28. About how many people live in your city?",
                    "29. Is your country of citizenship different from the country where you were born?",
                    "30. Are you an ethnic minority in your country?",
                    "31. Is English your second language?",
                    "35. What's your marital status?",
                    "36. How many children do you have? By children, we mean any biological or adopted children",
                    "37. Do you financially support any dependents, older relatives or other family members?",
                    "38. Do you think you have enough savings to survive for 3 months if you lost your job?",
                    "39. How much debt does your household have? [Car loans]",
                    "39. How much debt does your household have? [Credit Cards]",
                    "39. How much debt does your household have? [Medical Loans]",
                    "39. How much debt does your household have? [Mortgages]",

```

```
"39. How much debt does your household have? [Payday Loans]",
"39. How much debt does your household have? [Student Loans]",
"39. How much debt does your household have? [Other]",
"46. About how many minutes does it take you to get to work each d
"47. Have you served in your country's military before?",
"48. Do you currently receive disability benefits from your govern
"49. Do you have high speed internet at your home?"]
```

```
survey_df.drop(columns_to_remove, axis=1, inplace = True)
```

```
survey_df.shape
```

```
(18126, 43)
```

We have removed 20 columns in our dataset.

## Sample

We can take a small sample of our dataset to get an idea of what type of data we are working with.

```
survey_df.sample(5)
```

	1. What is your biggest reason for learning to code?	2. What methods have you used to learn about coding? Please select all that apply.	3. Which online learning resources have you found helpful? Please select all that apply.	4. If you have attended in-person coding-related events before, which ones have you found helpful? Please select all that apply.	5. If you have listened to coding-related podcasts before, which ones have you found helpful? Please select all that apply.	6. If you have watched coding-related YouTube videos before, which channels have you found helpful? Please select all that apply.	7. About how many hours do you spend learning each week?	8. About how many months have you been programming?	9. As fr univers tuiti ab h mu mor have y spent learni to cc so far Dollar
353	As a hobby	Online resources, Online bootcamps	freeCodeCamp, Mozilla Developer Network (MDN),...	I haven't attended any in-person coding-relate...	I haven't listened to any podcasts	Coder, Coding Train, Dev Ed	10.0	28	
2444	As a hobby	Books, School Subject	freeCodeCamp	weekend bootcamps	I haven't listened to any podcasts	freeCodeCamp	1.5	4	
16087	To change careers	Online resources, Hackathons	freeCodeCamp, Mozilla Developer Network (MDN),...	I haven't attended any in-person coding-relate...	I haven't listened to any podcasts	CS Dojo, freeCodeCamp, Programming With Mosh	5.0	18	2
4645	To change careers	Online resources	freeCodeCamp, Udemy	NaN	NaN	NaN	20.0	1	5

	1. What is your biggest reason for learning to code?	2. What methods have you used to learn about coding? Please select all that apply.	3. Which online learning resources have you found helpful? Please select all that apply.	4. If you have attended in-person coding-related events before, which ones have you found helpful? Please select all that apply.	5. If you have listened to coding-related podcasts before, which ones have you found helpful? Please select all that apply.	6. If you have watched coding-related YouTube videos before, which channels have you found helpful? Please select all that apply.	7. About how many hours do you spend learning each week?	8. About how many months have you been programming?	9. As far as you know, how much do you think you will spend on your education in the next 12 months? Please select all that apply.
15335	To start your first career	Online resources, College	freeCodeCamp, Mozilla Developer Network (MDN),...	I haven't attended any in-person coding-relate...	I haven't listened to any podcasts	Programming With Mosh	20.0	12	\$0.00

## Checking for missing values

```
survey_df.isna()
```



18122	False	False	False	False	False	False	True	True	True	False
18123	False	False	False	False	False	False	False	False	False	False
18124	False	False	False	False	False	False	False	False	False	False
18125	False	False	False	False	False	False	False	False	False	False

18126 rows × 43 columns

```
survey_df.isna().sum().sort_values(ascending = False)
```

43. If you are working, how easy would it be for you to find a job with another employer with approximately the same income and fringe benefits you now have?

7212

42. If you are working, thinking about the next 12 months, how likely do you think it is that you will lose your job or be laid off?

7137

45. Please tell us how satisfied you are with each of these following aspects of your present job [Diverse and inclusive work environment]

5917

45. Please tell us how satisfied you are with each of these following aspects of your present job [Weekly workload]

5904

45. Please tell us how satisfied you are with each of these following aspects of your present job [Workplace/company culture]

5882

45. Please tell us how satisfied you are with each of these following aspects of your present job [Professional growth or leadership opportunities]

5877

45. Please tell us how satisfied you are with each of these following aspects of your present job [Job security]

5867

45. Please tell us how satisfied you are with each of these following aspects of your present job [Work-life balance]

5863

45. Please tell us how satisfied you are with each of these following aspects of your present job [Benefits]

5847

41. Before you got your last job, how many months did you spend looking for a job?

5774

45. Please tell us how satisfied you are with each of these following aspects of your present job [Earnings]

5769

27. If you are living in the US, which state do you currently live in?

4802

44. Do you consider yourself under-employed? (Under-employment means working a job that is below your education level.)

3105

22. About how much money did you earn last year from any job or employment (in US Dollars)?

2745

16. About how much money do you expect to earn per year at your first developer job (in US Dollars)?  
2375

9. Aside from university tuition, about how much money have you spent on learning to code so far (in US Dollars)?  
1973

4. If you have attended in-person coding-related events before, which ones have you found helpful? Please select all that apply.  
1531

5. If you have listened to coding-related podcasts before, which ones have you found helpful? Please select all that apply.  
1528

7. About how many hours do you spend learning each week?  
1523

8. About how many months have you been programming?  
1431

6. If you have watched coding-related YouTube videos before, which channels have you found helpful? Please select all that apply.  
1255

21. If you are currently working, which field do you work in?  
1242

40. How long have you been working in your current job?  
1183

33. If you attended a university, what did you study?  
1161

23. How old are you?  
1023

18. After the pandemic, how many days would you ideally like to work from home versus in an office each week?  
938

25. With which of these groups do you primarily identify?  
935

32. What is the highest degree or level of school you have completed?  
921

11. If you are already employed as a developer, is this your first software development job?  
904

26. Which part of the world do you live in?  
892

34. Are you currently a student attending or enrolled in regular school, that is in an elementary school, a middle school, a high school, a college, or a graduate school?  
852

20. Regarding employment status, are you currently....  
829

19. Are you willing to relocate for a job?  
760

3. Which online learning resources have you found helpful? Please select all that apply.  
736

14. Which of these careers are you interested in?

735

24. Which of the following best represents how you think of yourself?

664

13. If you are interested in a software development career, would you prefer to...

657

12. If you are NOT already a developer, are you interested in a software development career?

589

15. When do you plan to start applying for developer jobs?

573

17. Please select up to 3 reasons why you are interested in a software development career

538

10. Are you already employed in a software development job?

453

2. What methods have you used to learn about coding? Please select all that apply.

266

1. What is your biggest reason for learning to code?

135

dtype: int64

There are many methods that can be used to handle missing values from datasets, with 'dropna' and 'fillna' being the most common. For now, we will leave the dataset as it is.

## Changing column names and creating new Dataframe

The names of our columns are fairly long, so it would be easier to rename them with abbreviated names. To keep track of the questions, we will create a pandas dataframe which consists of the new column names and the questions associated with them i.e. the original column names.

First, we will look at what columns we have.

```
survey_df.columns
```

```
Index(['1. What is your biggest reason for learning to code?',  
      '2. What methods have you used to learn about coding? Please select all that  
      apply.',  
      '3. Which online learning resources have you found helpful? Please select all  
      that apply.',  
      '4. If you have attended in-person coding-related events before, which ones have  
      you found helpful? Please select all that apply.',  
      '5. If you have listened to coding-related podcasts before, which ones have you  
      found helpful? Please select all that apply.',  
      '6. If you have watched coding-related YouTube videos before, which channels  
      have you found helpful? Please select all that apply.',  
      '7. About how many hours do you spend learning each week?',  
      '8. About how many months have you been programming?',  
      '9. Aside from university tuition, about how much money have you spent on  
      learning to code so far (in US Dollars)?',
```

'10. Are you already employed in a software development job?',  
'11. If you are already employed as a developer, is this your first software development job? ',  
'12. If you are NOT already a developer, are you interested in a software development career?',  
'13. If you are interested in a software development career, would you prefer to...',  
'14. Which of these careers are you interested in?',  
'15. When do you plan to start applying for developer jobs?',  
'16. About how much money do you expect to earn per year at your first developer job (in US Dollars)? ',  
'17. Please select up to 3 reasons why you are interested in a software development career',  
'18. After the pandemic, how many days would you ideally like to work from home versus in an office each week?',  
'19. Are you willing to relocate for a job?',  
'20. Regarding employment status, are you currently.... ',  
'21. If you are currently working, which field do you work in?',  
'22. About how much money did you earn last year from any job or employment (in US Dollars)? ',  
'23. How old are you?',  
'24. Which of the following best represents how you think of yourself?',  
'25. With which of these groups do you primarily identify?',  
'26. Which part of the world do you live in?',  
'27. If you are living in the US, which state do you currently live in? ',  
'32. What is the highest degree or level of school you have completed?',  
'33. If you attended a university, what did you study?',  
'34. Are you currently a student attending or enrolled in regular school, that is in an elementary school, a middle school, a high school, a college, or a graduate school?',  
'40. How long have you been working in your current job?',  
'41. Before you got your last job, how many months did you spend looking for a job?',  
'42. If you are working, thinking about the next 12 months, how likely do you think it is that you will lose your job or be laid off? ',  
'43. If you are working, how easy would it be for you to find a job with another employer with approximately the same income and fringe benefits you now have? ',  
'44. Do you consider yourself under-employed? (Under-employment means working a job that is below your education level.)',  
'45. Please tell us how satisfied you are with each of these following aspects of your present job [Earnings]',  
'45. Please tell us how satisfied you are with each of these following aspects of your present job [Benefits]',  
'45. Please tell us how satisfied you are with each of these following aspects of your present job [Job security]',  
'45. Please tell us how satisfied you are with each of these following aspects of your present job [Work-life balance]',  
'45. Please tell us how satisfied you are with each of these following aspects of your present job [Professional growth or leadership opportunities]',

```
'45. Please tell us how satisfied you are with each of these following aspects  
of your present job [Workplace/company culture]',  
'45. Please tell us how satisfied you are with each of these following aspects  
of your present job [Diverse and inclusive work environment]',  
'45. Please tell us how satisfied you are with each of these following aspects  
of your present job [Weekly workload]'],  
dtype='object')
```

Now we will create two lists which contain our new column names and old column names.

```
new_column_names_list = [ "BiggestReasonToCode",  
                           "MethodsUsedToLearn",  
                           "OnlineResourcesUsed",  
                           "UsefulCodingEvents",  
                           "UsefulCodingPodcasts",  
                           "UsefulCodingVideos",  
                           "HoursSpentLearningPerWeek",  
                           "MonthsSpentProgramming",  
                           "MoneySpentLearning",  
                           "AlreadyEmployed",  
                           "FirstSoftwareJob",  
                           "InterestedInSoftware",  
                           "PreferredCareerPath",  
                           "CareersInterestedIn",  
                           "WhenApplyingForJobs",  
                           "ExpectedSalary",  
                           "ReasonsForSoftwareInterest",  
                           "WorkFromHomeVsOffice",  
                           "WillingToRelocate",  
                           "EmploymentStatus",  
                           "CurrentFieldOfWork",  
                           "MoneyMadeLastYear",  
                           "Age",  
                           "Gender",  
                           "Ethnicity",  
                           "Region",  
                           "State",  
                           "EducationLevel",  
                           "UniversityDegree",  
                           "CurrentStudent",  
                           "TimeSpentInCurrentRole",  
                           "MonthsSpentLookingForJob",  
                           "LikelyToLoseJob",  
                           "FindNewJobEasiness",  
                           "ConsideredUnderEmployed",  
                           "EarningSatisfaction",  
                           "BenefitsSatisfaction",  
                           "JobSecuritySatisfaction",  
                           "WorkLifeBalanceSatisfaction",  
                           "GrowthOpportunitySatisfaction",  
                           "WorkCultureSatisfaction",
```

```
"WorkDiversitySatisfaction",  
"WorkloadSatisfaction"]
```

```
len(new_column_names_list)
```

43

[illegible]

```
len(questions_list)
```

43

We will pair up our column names and questions with the `zip()` function.

```
zipped = list(zip(new_column_names_list, questions_list))
```

```
zipped
```

```
[('BiggestReasonToCode',  
  '1. What is your biggest reason for learning to code?'),  
( 'MethodsUsedToLearn',  
  '2. What methods have you used to learn about coding? Please select all that  
apply.'),  
( 'OnlineResourcesUsed',  
  '3. Which online learning resources have you found helpful? Please select all that  
apply.'),  
( 'UsefulCodingEvents',  
  '4. If you have attended in-person coding-related events before, which ones have you  
found helpful? Please select all that apply.'),  
( 'UsefulCodingPodcasts',  
  '5. If you have listened to coding-related podcasts before, which ones have you found  
helpful? Please select all that apply.'),  
( 'UsefulCodingVideos',  
  '6. If you have watched coding-related YouTube videos before, which channels have you  
found helpful? Please select all that apply.'),  
( 'HoursSpentLearningPerWeek',  
  '7. About how many hours do you spend learning each week?'),  
( 'MonthsSpentProgramming',  
  '8. About how many months have you been programming?'),  
( 'MoneySpentLearning',  
  '9. Aside from university tuition, about how much money have you spent on learning to  
code so far (in US Dollars)?'),  
( 'AlreadyEmployed',  
  '10. Are you already employed in a software development job?'),  
( 'FirstSoftwareJob',  
  '11. If you are already employed as a developer, is this your first software  
development job? '),  
( 'InterestedInSoftware',  
  '12. If you are NOT already a developer, are you interested in a software development  
career?'),  
( 'PreferredCareerPath',  
  '13. If you are interested in a software development career, would you prefer  
to...'),  
( 'CareersInterestedIn', '14. Which of these careers are you interested in?'),  
( 'WhenApplyingForJobs',  
  '15. When do you plan to start applying for developer jobs?'),  
( 'ExpectedSalary',
```

'16. About how much money do you expect to earn per year at your first developer job (in US Dollars)? '),  
( 'ReasonsForSoftwareInterest',  
'17. Please select up to 3 reasons why you are interested in a software development career'),  
( 'WorkFromHomeVsOffice',  
'18. After the pandemic, how many days would you ideally like to work from home versus in an office each week?'),  
( 'WillingToRelocate', '19. Are you willing to relocate for a job?'),  
( 'EmploymentStatus',  
'20. Regarding employment status, are you currently.... '),  
( 'CurrentFieldOfWork',  
'21. If you are currently working, which field do you work in?'),  
( 'MoneyMadeLastYear',  
'22. About how much money did you earn last year from any job or employment (in US Dollars)? '),  
( 'Age', '23. How old are you?'),  
( 'Gender',  
'24. Which of the following best represents how you think of yourself?'),  
( 'Ethnicity', '25. With which of these groups do you primarily identify?'),  
( 'Region', '26. Which part of the world do you live in?'),  
( 'State',  
'27. If you are living in the US, which state do you currently live in? '),  
( 'EducationLevel',  
'32. What is the highest degree or level of school you have completed?'),  
( 'UniversityDegree', '33. If you attended a university, what did you study?'),  
( 'CurrentStudent',  
'34. Are you currently a student attending or enrolled in regular school, that is in an elementary school, a middle school, a high school, a college, or a graduate school?'),  
( 'TimeSpentInCurrentRole',  
'40. How long have you been working in your current job?'),  
( 'MonthsSpentLookingForJob',  
'41. Before you got your last job, how many months did you spend looking for a job?'),  
( 'LikelyToLoseJob',  
'42. If you are working, thinking about the next 12 months, how likely do you think it is that you will lose your job or be laid off? '),  
( 'FindNewJobEasiness',  
'43. If you are working, how easy would it be for you to find a job with another employer with approximately the same income and fringe benefits you now have? '),  
( 'ConsideredUnderEmployed',  
'44. Do you consider yourself under-employed? (Under-employment means working a job that is below your education level.)'),  
( 'EarningSatisfaction',  
'45. Please tell us how satisfied you are with each of these following aspects of your present job [Earnings]'),  
( 'BenefitsSatisfaction',  
'45. Please tell us how satisfied you are with each of these following aspects of



```

your present job [Benefits]'),
('JobSecuritySatisfaction',
'45. Please tell us how satisfied you are with each of these following aspects of
your present job [Job security]'),
('WorkLifeBalanceSatisfaction',
'45. Please tell us how satisfied you are with each of these following aspects of
your present job [Work-life balance]'),
('GrowthOpportunitySatisfaction',
'45. Please tell us how satisfied you are with each of these following aspects of
your present job [Professional growth or leadership opportunities]'),
('WorkCultureSatisfaction',
'45. Please tell us how satisfied you are with each of these following aspects of
your present job [Workplace/company culture]'),
('WorkDiversitySatisfaction',
'45. Please tell us how satisfied you are with each of these following aspects of
your present job [Diverse and inclusive work environment]'),
('WorkloadSatisfaction',
'45. Please tell us how satisfied you are with each of these following aspects of
your present job [Weekly workload]')])

```

We now have what we need to create our dataframe.

```
questions_df = pd.DataFrame(zippped, columns = ['Column', 'QuestionText'])
```

```
questions_df
```

	Column	QuestionText
0	BiggestReasonToCode	1. What is your biggest reason for learning to...
1	MethodsUsedToLearn	2. What methods have you used to learn about c...
2	OnlineResourcesUsed	3. Which online learning resources have you fo...
3	UsefulCodingEvents	4. If you have attended in-person coding-relat...
4	UsefulCodingPodcasts	5. If you have listened to coding-related podc...
5	UsefulCodingVideos	6. If you have watched coding-related YouTube ...
6	HoursSpentLearningPerWeek	7. About how many hours do you spend learning ...
7	MonthsSpentProgramming	8. About how many months have you been program...
8	MoneySpentLearning	9. Aside from university tuition, about how mu...
9	AlreadyEmployed	10. Are you already employed in a software dev...
10	FirstSoftwareJob	11. If you are already employed as a developer...
11	InterestedInSoftware	12. If you are NOT already a developer, are yo...
12	PreferredCareerPath	13. If you are interested in a software develo...
13	CareersInterestedIn	14. Which of these careers are you interested in?
14	WhenApplyingForJobs	15. When do you plan to start applying for dev...
15	ExpectedSalary	16. About how much money do you expect to earn...
16	ReasonsForSoftwareInterest	17. Please select up to 3 reasons why you are ...
17	WorkFromHomeVsOffice	18. After the pandemic, how many days would yo...

	Column	QuestionText
18	WillingToRelocate	19. Are you willing to relocate for a job?
19	EmploymentStatus	20. Regarding employment status, are you curre...
20	CurrentFieldOfWork	21. If you are currently working, which field ...
21	MoneyMadeLastYear	22. About how much money did you earn last yea...
22	Age	23. How old are you?
23	Gender	24. Which of the following best represents how...
24	Ethnicity	25. With which of these groups do you primaril...
25	Region	26. Which part of the world do you live in?
26	State	27. If you are living in the US, which state d...
27	EducationLevel	32. What is the highest degree or level of sch...
28	UniversityDegree	33. If you attended a university, what did you...
29	CurrentStudent	34. Are you currently a student attending or e...
30	TimeSpentInCurrentRole	40. How long have you been working in your cur...
31	MonthsSpentLookingForJob	41. Before you got your last job, how many mon...
32	LikelyToLoseJob	42. If you are working, thinking about the nex...
33	FindNewJobEasiness	43. If you are working, how easy would it be f...
34	ConsideredUnderEmployed	44. Do you consider yourself under-employed? (...)
35	EarningSatisfaction	45. Please tell us how satisfied you are with ...
36	BenefitsSatisfaction	45. Please tell us how satisfied you are with ...
37	JobSecuritySatisfaction	45. Please tell us how satisfied you are with ...
38	WorkLifeBalanceSatisfaction	45. Please tell us how satisfied you are with ...
39	GrowthOpportunitySatisfaction	45. Please tell us how satisfied you are with ...
40	WorkCultureSatisfaction	45. Please tell us how satisfied you are with ...
41	WorkDiversitySatisfaction	45. Please tell us how satisfied you are with ...
42	WorkloadSatisfaction	45. Please tell us how satisfied you are with ...

We can simplify this further by converting our Dataframe into a Pandas Series, with 'Column' being the index. The reason we are doing this is we want to use this Dataframe to access the questions associated with our dataset, and we only have one column so storing our information as a Dataframe is unnecessary.

```
questions = questions_df.set_index('Column').QuestionText
```

```
questions
```

Column

BiggestReasonToCode	1. What is your biggest reason for learning to...
MethodsUsedToLearn	2. What methods have you used to learn about c...
OnlineResourcesUsed	3. Which online learning resources have you fo...
UsefulCodingEvents	4. If you have attended in-person coding-relat...
UsefulCodingPodcasts	5. If you have listened to coding-related podc...
UsefulCodingVideos	6. If you have watched coding-related YouTube ...
HoursSpentLearningPerWeek	7. About how many hours do you spend learning ...

MonthsSpentProgramming	8. About how many months have you been program...
MoneySpentLearning	9. Aside from university tuition, about how mu...
AlreadyEmployed	10. Are you already employed in a software dev...
FirstSoftwareJob	11. If you are already employed as a developer...
InterestedInSoftware	12. If you are NOT already a developer, are yo...
PreferredCareerPath	13. If you are interested in a software develo...
CareersInterestedIn	14. Which of these careers are you interested in?
WhenApplyingForJobs	15. When do you plan to start applying for dev...
ExpectedSalary	16. About how much money do you expect to earn...
ReasonsForSoftwareInterest	17. Please select up to 3 reasons why you are ...
WorkFromHomeVsOffice	18. After the pandemic, how many days would yo...
WillingToRelocate	19. Are you willing to relocate for a job?
EmploymentStatus	20. Regarding employment status, are you curre...
CurrentFieldOfWork	21. If you are currently working, which field ...
MoneyMadeLastYear	22. About how much money did you earn last yea...
Age	23. How old are you?
Gender	24. Which of the following best represents how...
Ethnicity	25. With which of these groups do you primaril...
Region	26. Which part of the world do you live in?
State	27. If you are living in the US, which state d...
EducationLevel	32. What is the highest degree or level of sch...
UniversityDegree	33. If you attended a university, what did you...
CurrentStudent	34. Are you currently a student attending or e...
TimeSpentInCurrentRole	40. How long have you been working in your cur...
MonthsSpentLookingForJob	41. Before you got your last job, how many mon...
LikelyToLoseJob	42. If you are working, thinking about the nex...
FindNewJobEasiness	43. If you are working, how easy would it be f...
ConsideredUnderEmployed	44. Do you consider yourself under-employed? (...)
EarningSatisfaction	45. Please tell us how satisfied you are with ...
BenefitsSatisfaction	45. Please tell us how satisfied you are with ...
JobSecuritySatisfaction	45. Please tell us how satisfied you are with ...
WorkLifeBalanceSatisfaction	45. Please tell us how satisfied you are with ...
GrowthOpportunitySatisfaction	45. Please tell us how satisfied you are with ...
WorkCultureSatisfaction	45. Please tell us how satisfied you are with ...
WorkDiversitySatisfaction	45. Please tell us how satisfied you are with ...
WorkloadSatisfaction	45. Please tell us how satisfied you are with ...

Name: QuestionText, dtype: object

```
questions['BiggestReasonToCode']
```

'1. What is your biggest reason for learning to code?'

Finally, we want to rename our columns with the simplified names.

```
survey_df.columns = new_column_names_list
```

```
survey_df
```

	BiggestReasonToCode	MethodsUsedToLearn	OnlineResourcesUsed	UsefulCodingEvents	UsefulCodingPodcasts
0	To succeed in current career	Online resources, Books, In-person bootcamps, ...	freeCodeCamp, Mozilla Developer Network (MDN),...	conferences, workshops, Meetup.com events	The Changelog
1	To change careers	Online resources, Books, Online bootcamps	freeCodeCamp, Mozilla Developer Network (MDN),...	I haven't attended any in-person coding-relate...	The Changelog, Code Newbie Podcast
2	To change careers	Online resources, Books, Hackathons, Meetup.co...	freeCodeCamp, Mozilla Developer Network (MDN),...	Meetup.com events	I haven't listened to any podcasts
3	As a hobby	Online resources, Books	freeCodeCamp, Mozilla Developer Network (MDN),...	I haven't attended any in-person coding-relate...	Darknet Diaries, Real Python Podcast
4	To start your first career	Online resources, Books, Online bootcamps	freeCodeCamp, Stack Overflow, Coursera, Udemy	I haven't attended any in-person coding-relate...	Talk Python to Me
...	...	...	...	...	...
18121	To change careers	Online resources, Books, In-person bootcamps	Stack Overflow, EdX, Khan Academy, Udacity, Ci...	conferences, workshops	Le code a changé (RadioFrance app)
18122	To change careers	Online resources, In-person bootcamps	freeCodeCamp, Mozilla Developer Network (MDN),...	I haven't attended any in-person coding-relate...	I haven't listened to any podcasts
18123	To start your first career	Online resources, Books	freeCodeCamp, Mozilla Developer Network (MDN),...	I haven't attended any in-person coding-relate...	I haven't listened to any podcasts
18124	To change careers	Online resources, Books, In-person bootcamps, ...	freeCodeCamp, Stack Overflow, Codecademy, Udem...	I haven't attended any in-person coding-relate...	I haven't listened to any podcasts
18125	To succeed in current career	Online resources, Books	freeCodeCamp, Mozilla Developer Network (MDN),...	I haven't attended any in-person coding-relate...	I haven't listened to any podcasts

18126 rows × 43 columns

```
survey_df.columns
```

```
Index(['BiggestReasonToCode', 'MethodsUsedToLearn', 'OnlineResourcesUsed',
      'UsefulCodingEvents', 'UsefulCodingPodcasts', 'UsefulCodingVideos',
      'HoursSpentLearningPerWeek', 'MonthsSpentProgramming',
      'MoneySpentLearning', 'AlreadyEmployed', 'FirstSoftwareJob',
      'InterestedInSoftware', 'PreferredCareerPath', 'CareersInterestedIn',
      'WhenApplyingForJobs', 'ExpectedSalary', 'ReasonsForSoftwareInterest',
      'WorkFromHomeVsOffice', 'WillingToRelocate', 'EmploymentStatus',
      'CurrentFieldOfWork', 'MoneyMadeLastYear', 'Age', 'Gender', 'Ethnicity',
      'Region', 'State', 'EducationLevel', 'UniversityDegree',
      'CurrentStudent', 'TimeSpentInCurrentRole', 'MonthsSpentLookingForJob',
      'LikelyToLoseJob', 'FindNewJobEasiness', 'ConsideredUnderEmployed',
      'EarningSatisfaction', 'BenefitsSatisfaction',
      'JobSecuritySatisfaction', 'WorkLifeBalanceSatisfaction',
      'GrowthOpportunitySatisfaction', 'WorkCultureSatisfaction',
      'WorkDiversitySatisfaction', 'WorkloadSatisfaction'],
      dtype='object')
```

This will make it a lot easier when we want to access columns in our dataset moving forward, and we will still be able to look at the bank of questions associated with each column when necessary.

## Modifying the data type of our columns

We can use the `.info()` method to gather some basic information about our dataset.

```
survey_df.info()
```

```
<class 'pandas.core.frame.DataFrame'>
```

```
RangeIndex: 18126 entries, 0 to 18125
```

```
Data columns (total 43 columns):
```

#	Column	Non-Null Count	Dtype
0	BiggestReasonToCode	17991 non-null	object
1	MethodsUsedToLearn	17860 non-null	object
2	OnlineResourcesUsed	17390 non-null	object
3	UsefulCodingEvents	16595 non-null	object
4	UsefulCodingPodcasts	16598 non-null	object
5	UsefulCodingVideos	16871 non-null	object
6	HoursSpentLearningPerWeek	16603 non-null	float64
7	MonthsSpentProgramming	16695 non-null	object
8	MoneySpentLearning	16153 non-null	object
9	AlreadyEmployed	17673 non-null	object
10	FirstSoftwareJob	17222 non-null	object
11	InterestedInSoftware	17537 non-null	object
12	PreferredCareerPath	17469 non-null	object
13	CareersInterestedIn	17391 non-null	object
14	WhenApplyingForJobs	17553 non-null	object
15	ExpectedSalary	15751 non-null	object
16	ReasonsForSoftwareInterest	17588 non-null	object
17	WorkFromHomeVsOffice	17188 non-null	object
18	WillingToRelocate	17366 non-null	object
19	EmploymentStatus	17297 non-null	object
20	CurrentFieldOfWork	16884 non-null	object
21	MoneyMadeLastYear	15381 non-null	object
22	Age	17103 non-null	float64
23	Gender	17462 non-null	object
24	Ethnicity	17191 non-null	object
25	Region	17234 non-null	object
26	State	13324 non-null	object
27	EducationLevel	17205 non-null	object
28	UniversityDegree	16965 non-null	object
29	CurrentStudent	17274 non-null	object

30	TimeSpentInCurrentRole	16943	non-null	object
31	MonthsSpentLookingForJob	12352	non-null	float64
32	LikelyToLoseJob	10989	non-null	float64
33	FindNewJobEasiness	10914	non-null	float64
34	ConsideredUnderEmployed	15021	non-null	object
35	EarningSatisfaction	12357	non-null	object
36	BenefitsSatisfaction	12279	non-null	object
37	JobSecuritySatisfaction	12259	non-null	object
38	WorkLifeBalanceSatisfaction	12263	non-null	object
39	GrowthOpportunitySatisfaction	12249	non-null	object
40	WorkCultureSatisfaction	12244	non-null	object
41	WorkDiversitySatisfaction	12209	non-null	object
42	WorkloadSatisfaction	12222	non-null	object

dtypes: float64(5), object(38)

memory usage: 5.9+ MB

The majority of the columns in the dataset have the datatype 'Object' as they contain values of different types or have NaN values. Some columns such as MonthsSpentProgramming and MoneySpentLearning are Object datatypes despite having mainly numeric values. We can use the .to\_numeric() method to fix this for when we want to use these columns in the future.

```
survey_df['MonthsSpentProgramming'] = pd.to_numeric(survey_df.MonthsSpentProgramming, errors='coerce')
survey_df['MoneySpentLearning'] = pd.to_numeric(survey_df.MoneySpentLearning, errors='coerce')
```

```
survey_df.info()
```

```
<class 'pandas.core.frame.DataFrame'>
```

```
RangeIndex: 18126 entries, 0 to 18125
```

```
Data columns (total 43 columns):
```

#	Column	Non-Null Count	Dtype
0	BiggestReasonToCode	17991 non-null	object
1	MethodsUsedToLearn	17860 non-null	object
2	OnlineResourcesUsed	17390 non-null	object
3	UsefulCodingEvents	16595 non-null	object
4	UsefulCodingPodcasts	16598 non-null	object
5	UsefulCodingVideos	16871 non-null	object
6	HoursSpentLearningPerWeek	16603 non-null	float64
7	MonthsSpentProgramming	16695 non-null	float64
8	MoneySpentLearning	16153 non-null	float64
9	AlreadyEmployed	17673 non-null	object
10	FirstSoftwareJob	17222 non-null	object
11	InterestedInSoftware	17537 non-null	object
12	PreferredCareerPath	17469 non-null	object

13	CareersInterestedIn	17391	non-null	object
14	WhenApplyingForJobs	17553	non-null	object
15	ExpectedSalary	15751	non-null	object
16	ReasonsForSoftwareInterest	17588	non-null	object
17	WorkFromHomeVsOffice	17188	non-null	object
18	WillingToRelocate	17366	non-null	object
19	EmploymentStatus	17297	non-null	object
20	CurrentFieldOfWork	16884	non-null	object
21	MoneyMadeLastYear	15381	non-null	object
22	Age	17103	non-null	float64
23	Gender	17462	non-null	object
24	Ethnicity	17191	non-null	object
25	Region	17234	non-null	object
26	State	13324	non-null	object
27	EducationLevel	17205	non-null	object
28	UniversityDegree	16965	non-null	object
29	CurrentStudent	17274	non-null	object
30	TimeSpentInCurrentRole	16943	non-null	object
31	MonthsSpentLookingForJob	12352	non-null	float64
32	LikelyToLoseJob	10989	non-null	float64
33	FindNewJobEasiness	10914	non-null	float64
34	ConsideredUnderEmployed	15021	non-null	object
35	EarningSatisfaction	12357	non-null	object
36	BenefitsSatisfaction	12279	non-null	object
37	JobSecuritySatisfaction	12259	non-null	object
38	WorkLifeBalanceSatisfaction	12263	non-null	object
39	GrowthOpportunitySatisfaction	12249	non-null	object
40	WorkCultureSatisfaction	12244	non-null	object
41	WorkDiversitySatisfaction	12209	non-null	object
42	WorkloadSatisfaction	12222	non-null	object

dtypes: float64(7), object(36)

memory usage: 5.9+ MB

We have now changed the datatype of these two columns to float64.

## Addressing anomalies in the dataset

For the next step of preprocessing and cleaning our data for analysis, we want to take a closer look at the numeric columns in our dataset and check the values are sensible.

To start, we will use the `.describe()` function to look at some basic statistics about our data.

```
survey_df.describe()
```

HoursSpentLearningPerWeek	MonthsSpentProgramming	MoneySpentLearning	Age	MonthsSpentLook
---------------------------	------------------------	--------------------	-----	-----------------

	HoursSpentLearningPerWeek	MonthsSpentProgramming	MoneySpentLearning	Age	MonthsSpentLookingForJob
count	16603.000000	1.669500e+04	1.615300e+04	17103.000000	1.230000e+06
mean	12.654852	5.989817e+36	6.190800e+117	27.107420	5.640000e+05
std	13.583577	7.739391e+38	7.868164e+119	10.145221	6.270000e+05
min	0.000000	0.000000e+00	0.000000e+00	6.000000	0.000000e+00
25%	4.000000	1.000000e+00	0.000000e+00	20.000000	0.000000e+00
50%	8.000000	4.000000e+00	0.000000e+00	25.000000	1.000000e+00
75%	19.000000	1.200000e+01	6.700000e+01	32.000000	4.000000e+00
max	150.000000	1.000000e+41	1.000000e+122	120.000000	6.900000e+05

There seems to be an issue with some of the numerical values in our dataset. To start with, the maximum age is 120, which is not likely to be true. Secondly, the max value for number of hours spent learning each week is 150, which again seems unlikely. Finally, the average and max number of months spent looking for a job seems excessively high as well. We will fix these one at a time.

We will first look at the Age column. We will use the `.drop()` method to remove any rows where the age is greater than 100.

```
survey_df.Age.sort_values(ascending=False).head(5)
```

```
14136    120.0
13561    100.0
7548     100.0
84       100.0
2490     99.0
Name: Age, dtype: float64
```

```
survey_df.drop(survey_df[survey_df.Age > 100].index, inplace=True)
```

```
survey_df.Age.sort_values(ascending=False).head(5)
```

```
84       100.0
7548     100.0
13561    100.0
2490     99.0
10822    95.0
Name: Age, dtype: float64
```

Next up; the HoursSpentLearningPerWeek column. To fix this, we will drop any rows where the value is > 120.

```
survey_df.HoursSpentLearningPerWeek.sort_values(ascending=False).head(5)
```

```
11625    150.0
15564    150.0
13805    140.0
3841     133.0
12492    128.0
Name: HoursSpentLearningPerWeek, dtype: float64
```



```
survey_df.drop(survey_df[survey_df.HoursSpentLearningPerWeek > 120].index, inplace=True)
```

```
survey_df.HoursSpentLearningPerWeek.sort_values(ascending=False).head(5)
```

```
3522      120.0
13416     120.0
12097     120.0
4437      120.0
13666     120.0
Name: HoursSpentLearningPerWeek, dtype: float64
```

Finally, we can fix the MonthsSpentLookingForJob. We will use the same methodology as before to remove excessively high values (higher than 24 months or 2 years).

```
len(survey_df[survey_df.MonthsSpentLookingForJob > 24])
```

```
181
```

```
survey_df.MonthsSpentLookingForJob.sort_values(ascending=False).head(10)
```

```
2509      6.969700e+21
392       8.080808e+15
1301      1.000000e+13
2490      2.333333e+06
1287      9.990000e+02
18003     5.510000e+02
12076     2.810000e+02
14207     2.400000e+02
16498     2.400000e+02
12343     2.080000e+02
Name: MonthsSpentLookingForJob, dtype: float64
```

```
survey_df.drop(survey_df[survey_df.MonthsSpentLookingForJob > 24].index, inplace=True)
```

```
survey_df.MonthsSpentLookingForJob.sort_values(ascending=False).head(10)
```

```
5397      24.0
7887      24.0
15071     24.0
3679      24.0
5420      24.0
2041      24.0
15240     24.0
16751     24.0
9042      24.0
7749      24.0
Name: MonthsSpentLookingForJob, dtype: float64
```

```
survey_df.describe()
```

	HoursSpentLearningPerWeek	MonthsSpentProgramming	MoneySpentLearning	Age	MonthsSpentLookingForWork
count	16420.000000	16514.000000	1.597700e+04	16916.000000	121
mean	12.600396	15.618836	5.903804e+02	27.030515	
std	13.381484	100.060505	2.032150e+04	10.081573	
min	0.000000	0.000000	0.000000e+00	6.000000	
25%	4.000000	1.000000	0.000000e+00	20.000000	
50%	8.000000	4.000000	0.000000e+00	24.000000	
75%	18.000000	12.000000	6.000000e+01	32.000000	
max	120.000000	9000.000000	2.500001e+06	100.000000	

We have now cleaned up the erroneous numerical entries in our dataset and we can now move on to the next stage of this project; Exploratory Analysis and Visualization.

```
import jovian
```

```
jovian.commit()
```

```
[jovian] Updating notebook "adammasongh/2021-new-coder-survey" on https://jovian.com
[jovian] Committed successfully! https://jovian.com/adammasongh/2021-new-coder-survey
'https://jovian.com/adammasongh/2021-new-coder-survey'
```

## Exploratory Analysis and Visualization

In this section we will focus on getting a better understanding of the demographics of the respondents, to help give us an idea of the type of people who are interested in coding. We will focus on 5 key areas:

1. The respondent's biggest reasons for learning how to code.
2. Their gender.
3. Their ethnicity.
4. What they studied.
5. The highest degree they have obtained.

Exploring these variables will help us understand how representative the survey is of the programming community on a global scale, although please do note that any survey will have at least some degree of selection bias.

Let's begin by importing `matplotlib.pyplot` and `seaborn`.

```
import seaborn as sns
import matplotlib
import matplotlib.pyplot as plt
%matplotlib inline
```

```
sns.set_style('darkgrid')
matplotlib.rcParams['font.size'] = 14
matplotlib.rcParams['figure.figsize'] = (9, 5)
matplotlib.rcParams['figure.facecolor'] = '#00000000'
```

## Biggest reason for learning how to code

```
questions['BiggestReasonToCode']
```

```
'1. What is your biggest reason for learning to code?'
```

```
survey_df.BiggestReasonToCode.unique()
```

```
array(['To succeed in current career', 'To change careers', 'As a hobby',
      'To start your first career',
      'To start a business or to freelance',
      'To create art or entertainment', nan, 'Good question',
      'Have passion in coding',
      'When I was younger I used to play games a lot and wanted to create a game of my
own ;)',
      'I learned to code once I landed a job in an IT company that selected me based
on logical reasoning assessment',
      'To make videogames', 'To meet school requirements',
      'I want to create something awesome', 'To prepare for my career',
      'Its a passion and career. I love to build stuff',
      'To re enter workforce after being a stay at home parent for 10+ years',
      'to change the world ! ', 'I am a CS major', 'it is fascinating',
      'I made two websites on Wix and have been maintaining them myself, which seems
to be requiring some amount of coding these days. I cannot afford to hire a
professional, so...',
      'To start your first career, As a hobby and To create art or entertainment.',
      'My teen time dream (to be a coder) must becomes true, and to escape from my
country',
      'To make the world a better place.',
      "I left the platform because an author of one of the articles was talking about
his sexuality rather than the topic at hand. I am not interested in learning about
people's sexualities. I found it offensive and never looked back.I didnt get a job in
my field, maybe due to racism or prejudice or grades",
      'Improve on what I already know',
      'To retrain my brain to learn new things and focus, and also to change careers',
      'To outlive my obsolete-ness in IT industry of 15 years without having to code a
single line ever...like really EVER.',
      'To become a creator, not performer', 'Curiosity',
      'for academic research projects',
      'Always Learning, always changing, 30yr developer here',
      "It's fun and feels like a calling-something I am ment to do",
      "it's my passion",
      'I fell into it. Was a working graphic designer first.',
```

'Secondary income',  
'Hobby, current career, prep for future career change',  
'Just to learn and because my father pushed me in this direction, and I'm glad he did. ',  
'to learn math and science', "It's a childhood dream. ",  
'To create awesome programs so that I can help people',  
'62 year old unemployed', 'To Practice before 42 school',  
'As a hobby, To create art or entertainment, To change careers',  
'Enrolled in Computer Science program for bachelors', 'I love it.',  
'To start your first career and hobby as it is more than a career!.',  
'To create my own website(s), to help with my thesis',  
'to help process data at work',  
'I just had a feeling and ended up loving it. ', 'It pays better',  
'Change careers where I can make more money',  
'To change careers; As a hobby',  
'I want to be able to switch between professions while doing things for fun',  
'To change a careers and do what I like',  
'I am really interested in the field of ML, AI and IoT. Also you can use your programming knowledge in various engineering fields.',  
'to improve skills', 'I am still deciding.',  
'To change the world in my own way ',  
'To grow into a different arena of my career.',  
'Brushing up old skills while learning new ones.',  
"To change the world. I'm legally blind, companies don't design with my disability in mind.",  
'i have always had a connection with computers and electronics, so coding is more of the glue that holds or brings all of my computer sciencee love together.',  
'Additional career - have different daytime job and then code on the side (for work, own pleasure, help other bec i want to develop somethin that would be beneficial for most)',  
'I like coding',  
'Just want to get a normal job and continue programming in shadows.',  
'To be versatile in career options/knowledge',  
'I will need to code for my future/dream job.',  
'both passion and career', 'personal challenge and maybe more',  
'Curiosity and to help students begin their own journeys.',  
"It provides the only oppurtunity I've found to get paid doing what I love. so for money and happiness",  
'To have knowledge that can assure me some type of job',  
"I'm passionate about CS and also currently studying it at college. Hence, I'm always eager to learn as much as possible about every aspect of this immense field. ",  
"It's fun ", 'to make robots',  
"I'm a college student trying to be more marketable",  
'Because I like it', 'I really enjoy learning about computers.',  
'I really like it. I took some programming courses yrs ago(2 decades) ',  
'biggest reason is to challenge myself, 2nd reason career change',  
'About to take PMP exam. This will be great knowledge for me if I am working software development projects (I believe). ',  
'To improve my mental health',

"I'm unemployed. I need to learn programming to get a new job",  
'To find a good paying job',  
'Para obtener lo conocimientos básicos para construir un proyecto web ',  
'To change my life', 'To supplement my scientific research career',  
'To restart my career.', 'To develop more my programming interest',  
'As a component to inform my future career (entrepreneurship)',  
'to learn a new thing.',

"For my first career, but overall, coding just excites me in a way that doesn't exist in any other market.",

'I enjoy learning new things and also like the idea of knowing how the things I use often function so that I can create/fix those things myself. ',

"Reasons 1, 4 & 6; I've always enjoyed coding (especially for artistic expression), but now I'm serious about honing my craft because I want a career that offers me more time to meander and muse. The freelance wave seems to flash that motif a lot.",

'as a 10y old and I also have made a site', 'create a side income',  
'I just like it', 'To become financially independent',  
'to be able to initiate and teach my kids early', 'To make money',  
'To implement my creativity',

"Very Much Passionate about how all things work, what's behind the scene? that's why and also to succeed in career and to start business or freelance :-)",

'All the above',  
'Everything except first option(To change Career);',  
'As side resource from college', 'I <3 coding',  
'Career branch "shift". I\'m already in IT, but I\'d rather do coding.',  
'Coding is life',

'it gives a power of writing your imagination, you can understand how a technology works(at least on the software side). and you can create anything ( most ) and can automate any task',

"it is my passion from childhood even though i'm in a different a career at present, i want to get back to my passion of coding and make it into a career.",

'to change careers or maybe to freelance',  
'to get a head start as a major in college',

'to understand better the cryptoassets world, and be a better cryptoinvestor; but also for fun and be a freelancer sounds not bad',

"To challenge myself to learn sth I used to think of as impossible for me (got a lot of sexist 'girls are no good in tech' crap growing up). And to improve my work-at-home prospects as a person with chronic illnessess.",

'Curiosity ', 'It was part of my passion ',  
'To access the massive number of opportunities in world of computer science.',  
'Unsure of where it will take me, but certain it will do me good!',  
'To generate a second income',

"at least 4 reasons from this card. this should've been multi radio input.",  
'I just love solving problems.', 'To challenge myself',  
'To learn a new skill', 'I wish to become a wizard',

"Looking to learn a new skill that's different from my others.",  
'when i grow up i want to have a computer job and i enjoy coding ',

'I would like to learn some basic coding skills before enrolling in university for a cybersecurity major',

'i feel like it',  
"As a hobby, to create art and games, to start my first career, and because it's cool",  
'To add to skillset', 'To prepare for university',  
'I want to play with code',  
'I was wanted to create some COOL stuff and certainly i m more passionate about computers than any other field, therefore ',  
'To expand current capabilities to facilitate career',  
'better quality of life',  
'To get a developer Job and run a business.', 'to make games',  
'Hobby and to change career or to build some aplication that I have in mind for some time now ;] ',  
'To build meaningful software and tools that creates value for society',  
"I always loved computer related skills and I'm a bit of a tech nerd, but circumstances werent' favorable for me to pursue.",  
'Cant explain just love it', 'To success in life, a remote life',  
'Originally for research in grad school but now I am thinking of changing careers because I hate playing with petri dishes and dissecting cute rats',  
'life long learner looking to expand my skill set ',  
'As a hobby and for future job opportunities.',  
'to make my own game ', 'For all!',  
'To develop a new skill that has potential to benefit me in the future',  
"I know it's an important skill to have.",  
'To create, develop and innovate technologies for next generation and to make a world better place to live.',  
'to create my own project', 'Earn money', 'Hobby and for carrear',  
'get knowledge',  
'its my first step to becoming a software engineer',  
'To change careers and is a hobby too.',  
'To be able to apply to well paid jobs', 'curious',  
'to create something that make me proud such as this that will be usefull for others',  
'Refresh my programming skills',  
'General interest in the subject, not sure if it will lead to a change in careers. I currently work in IT as it is.',  
'Como complemento al desarrollo profesional',  
'To develop my skill set', 'To succeed in future career', 'Money',  
'To create websites and apps and know how they are made',  
'To augment my current knowledge',  
'As a hobby to improve my career.', 'To earn money', 'Bored',  
'may need to immigrate from Hong Kong to other countries, because of China',  
'To Start my career as a Data Scientist and after sometime start a business or freelance work.',  
'To learn about virtual reality and mixed reality',  
'to just have computer science skills',  
'to fulfill a creative need',  
'To improve my skills, and be able to be prepared for new challenges. Eventually build something.',  
'Both hoping to get a dev job, and also eventually transition into video game

dev!!',  
    'Explore, I might want to start a career with it, but it might just be something cool to learn',  
    'To add another possibility and specificity in career choice.',  
    '2, 4, 5',  
    'I want to know if I would be interested in this type of career',  
    'To make my own software house',  
    'To enhance my entrepreneurial journey. This is the skill I need to conquer the world.',  
    'To prepare for a career while at Seniors in High School',  
    'to learn something new and creative while designing for others',  
    'To expand my knowledge and create new opportunities in my current career path',  
,  
    'Hobby and change Careers', 'to explore my interests',  
    'To create something, whether for fun, or for profit',  
    'want to achieve something',  
    'Both to change careers and develop better skills for current career. Also willing to see if the skills developed will eventually be handy in creation of art',  
    'For a future job',  
    'I started coding because I had lost my job after 10 years of hard labor. I was diving into substance abuse and my children were watching who i was becoming. I wanted my children to see a brighter future so I started coding. I did receive an associates degree from a community college. ',  
    'curiosity', 'Me llama mucho la atención aprender a codificar',  
    'To teach my grandkids. This is their world and I want them to be able to make a living in this crazy world.',  
    'I like coding and I want to be a coder',  
    'Supplement my college education', 'To enhance my current career',  
    'Improve in my current career and possibly switch to a dev job from a design/development engineering job',  
    'Try out how programming is and see if i like it and can turn it into a career :D',  
    'Hobby AND succeed in current career', 'To improve my current job',  
    'I love it, and bc it's the future", "Want to try if it's for me ",  
    'to earn money from home ',  
    'I have been humiliated by people who knows how to code',  
    'To change the world',  
    'learn coding again, using current libraries like ReactJS',  
    "As a hedge against what I'm currently studying (psychology) but also because I think it's something new I want to try",  
    'I have nothing else to do', 'To change careers, and as a hobby.',  
    'Honestly multiple of these including, to start a career and freelance!',  
    'to learn new skill', 'to change careers AND freelance',  
    'To explore different opportunities of tech', 'I love it',  
    'To provide solutions to problem and change the world',  
    'To help improve not only my own skills, but to better maintain our websites.',  
    'for pussy', 'To do cool stuff and help people',  
    'I wanna improve my skills so I can get into the career field in the future.',  
    'as both a hobby and a second source of income that can be done from home in

spare time.',

'Originally to change careers, start first "real" career. Currently refreshing skills to improve career, change jobs, and out of personal interest',

'I see coding as a skillset that will soon be needed to hold a job, so to keep up and stand out in my Accounting role, I want to learn to code.',

'to understand how computers work today',

'Both as a hobby and to start my first career',

'To create a media company called Rinzai Media',

'I am 13 years old and I want to become a programmer.',

'to get a university degree',

'Keep updated and learning new things. Hope it will help me in the future',

'just curious and maybe change career if I had the chance', 'work',

'I just like to do it , because rather than sitting at home doing nothing is worthless and a waste of time, so my brother suggested me to learn coding from freecodecamp.org ',

'To discover my yet undiscovered part of me',

'want to be ahead in school',

'a combination between hobby and change carrer',

'To earning a living doing something that keeps me sharp while supporting myself and my family',

'I need to fix a text color in WordPress',

'Interest in how apps work',

'improve skills for to raise my salary',

'As a potential backup career plan',

'to change careers for remote jobs',

"I am an IT Professional who don't know how to code - WHAT?!? My aim is to change this.",

'Enjoy challenge of learning something new/fell in love with coding in 2015 when studying EE. Coming back to learn and finish what I started. ',

"Don't know anything about coding/curious",

"I want to leave school as it's too boring and is only about memorization. I enjoy coding and I'm planning to learn coding until I get a job after I reach the legal age.",

"Because the goverment don't invest money in my carreer(philosophy teacher)",

'To prove that I can',

'I am 40 years old, I could never study and the code can give me a job that I can do in my country or in another. Thanks',

'to know how code works',

'To learn something new to improve my horizons.',

'Mix of hobby and potentially changing/modifying career ',

'To see if it interests me and also to try and bring ideas I have into fruition',

'to start a business and create art',

"I believe it's a skill everyone should develop nowadays.",

'all of the above',

'I own a start up medical device company. I taught myself CAD and now am teaching myself coding to augment the design process, specifically in regards to compliant mechanisms, auxetic materials, and "origami inspired" design elements.',

'to begin a new, meaningful career', 'A project in mind',



"Computers are everywhere: learning basic code is essential in today's world.",  
'so that in the future i defeat google', 'To bring revolution ',  
'To learn it as a tool and a language in order to understand the digital world  
in which I evolve',  
'broaden resume/learn',  
'I want to build things and I want to do something interesting',  
'To be useful ',  
"It's always have been great interest of coding in me , I'm already working in  
IT ",  
'to add a valuable skill to my repertoire',  
'To start my fourth career',  
'To start my coding career and as a hobby',  
'I initially chose to change careers, but it does mean much more to me than  
that. It could potentially mean more freedom and flexibility in my life by making it  
where I could freelance or not be stuck a jobs that I hate.',  
"Passion. I see solutions to issues, but don't have the tools to fix myself.  
Instead I have to direct a programmer to fix what I see.",  
'To acquire it as a skill',  
'I have a didactic idea I want to realize',  
'I mistakenly took it as my main college subject cause i thought its cool but  
reality sucks',  
'This was/is a Pandemic hobby; not sure if I am ever going to actually use it,  
but it makes me feel like I can accomplish at least one thing',  
'To solve problems',  
'To create some programs (price calculator and screenplay editor in Google Apps  
Script) to aid my current job (audio and video production)',  
'To be able to be a software engineer',  
'multiple reasons above. a hobby, and now potentially to start my first  
career!',  
'learn something new, be ready for future projects',  
'Improving my code skills',  
'I already know a fair amount but want a better foundation.',  
"Plan on making some small amount of income to save up for when I move out of my  
parent's house",  
'To learn a handy skill',  
'I'm a senior in highschool and plan on minoring in computer science but want  
some experience ',  
'as a hobby / to improve my cv to my future career.',  
'to learn valuable skills & so I can pursue any tech related endeavors that  
interest me in the future ',  
"Almost all the reason why i'll like to start coding and also Entrepreneur.",  
"To learn about it.it's not linked directly to my career. iam doing Mechanical  
engineering. but i believe it'll greatly help me.",  
'Find out if I like it',  
"I've wanted to become a programmer since i was a kid but i never really started  
learning it, and i just decided to learn it today.",  
'i just want to understand logic behind them.',  
'To change the world in a good way',  
'Its a mixture of Art & Entertainment and Starting a career.',

'To modify my career direction', 'Hobby and possibly a career',  
'To expand my skill set',  
'reskill/upskill to re-enter IT as a developer after long hiatus',  
'Work, and I like to create things', 'for knowledge',  
'To do stuff i want', "It's my dream and goal.", 'I need jobs.',  
'as a career but also as a hobby',  
'To do better at providing than my to be ex-husband.',  
'I am hoping to improve skills to grow in my current career, and also to start a business as a freelancer.',  
'I am not sure, I just wanted to code', 'to get fucking rich',  
'To change our country with the power of technology ',  
'Take advantage of Career Opportunity at work, Do freelance on the side.',  
'To have it as an extra skill',  
'To enhance my learning capabilities.00',  
'Make something of myself, find purpose, career.. change/improve life.. escape poverty, fight corruption and advance human rights. Make money.. It is just personal and deep and mixed up like that, lol.',  
'first career / career change / hobby / interest',  
'future-proof my career', 'All of the above', 'i live with it',  
'To start my first career AND make tge web easy and accessible for everyone and anyone.',  
'For money and the flexibility about different jobs',  
" it's profitable job",  
'Recently retired and wanting to do programming but work gets in the way.',  
'Professional skills - I am a programmer/developer',  
'to earn money through any skill. coding is just my first choice. ',  
'As a hobby and to write coding tutorials for tech blogs ',  
'I enjoy it', 'i love coding...',  
'To support my computational design skills.',  
'As a child, I want to have the resources required to get a headstart on my planned career.',  
'To have a remote job that pays well where I am not micromanaged',  
'Learn how things work (how computers, AI, electronics, math fit together)',  
'i have interest in coding and if i find it suting to me i plan to work in it in the future, i also wish to be able to create entertainment/art with coding for my own joy.',  
"they don't teach it in school",  
'To create a website and an app for my own church and businesses',  
'almost all of the above',  
'So I have a range of skills for a range of careers',  
'to survive, cus i had nothing better todo, its not expensive like pilot or doctor school.',  
'To contribute to the field of computer science and open source/dev community',  
'Freedom',  
'become a more critcal thinker, leverage the knowledge at current job and future roles as well.',  
'To create technology that'll help me in my tasks",  
'To have another career',  
'To support my family so that my wife can stay at home with our kids.',

'I'm hoping to start as a hobby, but if it grows to a passion, use it to shift my career.',  
'Homeless and need a job', 'To have to fall back on',  
'Need new meaning in life', 'To keep up with tech', 'to teach it',  
'There are multiple reasons for it want to start freelancing when I start with my college next year in which I will be hopefully studying business administration, it also more like a hobby for now but I am sure it going to help me in future when I actually go look for a job after my graduation ',  
'to get a jumpstart before university and get more women in coding!',  
'To give surprise to my fiancée by learning coding',  
'I wish to become a technical co-founder of a company one day. ',  
'it's my dream job', 'To make a needed program.',  
'To become someone who can impact the world and become dominant in the industry',  
'Something different I can do in retirement',  
'To create my own website or app',  
'Improve my skill with web dev knowledge',  
'Expand my current Software Engineering Professional Knowledge',  
'I want to be a developer in the future. That's my dream.',  
'I feel coding is an art and art can be used to join and create opportunities',  
'Want to create a non-profit website for national curriculum.',  
'to unlock the mystery of web creation and the wierd language it uses.',  
'To add personal value and for succeed in future works',  
'To enhance my current career. ', 'Personal growth',  
'More job opportunities ', 'To create wonderful tools',  
'to learn something new and current',  
'for hobby, to make games and get a career',  
'i liked it in high school',  
'Equally to change careers and to freelance; as well as to get ahead because I may go back to school for IT/Web Development.',  
'To see if coding is for me',  
'To improve my skills and be more flexible on my current job',  
'To enhance future career choices',  
'I've learned that no matter how many certifications, degrees, or bits of knowledge I have, I will never have job security, so I just keep collecting.',  
'My mom wants me to to get a good job',  
'I believe it has the power to bring great change. ',  
'To change careers, create art or entertainment and as a hobby',  
'both a hobby and school requirements+it could be a future career as well',  
'I like to learn new skills', 'to work on a school project',  
'back to work after long break',  
'As a hobby and possibly exploring career paths.',  
'Very intrigued by the digital world',  
'For fun and to be able to have a good skill to show on my resume',  
'To start earning something on the side, and maybe turn it into a career later on',  
'To return to a previous career by refreshing and building upon outdated skillset',  
'Curious',

'To become a dev at a minrcraft server and i wanna build stuff',  
'To excel at it and work towards building something that revolutionizes our way of living.',  
'Expand my knowledge base in my software testing career. ',  
'I like computers ',  
'Tôi cũng là một sinh viên khoa công nghệ thông tin nhưng công việc của tôi không phải ngồi hàng giờ viết code. Nhân dịp có thời gian rảnh tôi muốn học thêm điều gì đó...',  
'to create apps to provide me whis constant reinforcements of positive thoughts to puss me in the path of positive progressive growth',  
'I really like coding. But I am a newbie and I am looking forward to learn and make it my career for the future.',  
'I'm disabled and hope this help to stimulate my mind",  
'to expand my limits and i am always ready for extra ',  
'to create something new', 'To create a website',  
'To create a website for my YouTube channel',  
"Starting university (software engineering) next year and want to be ahead so i'm not stuck",  
"Gain more knowledge about coding because my uni doesn't teach much ",  
'I always wanted to learn to code, but felt intimidated and graduated in business administration instead',  
'To see if I can do it',  
'because I like to be innovative and I like to involve students in building their future',  
'just to know how to do it cause it could be usefull',  
'Both as a hobby and meeting school requirements',  
'para crear cosas que el ser humano necesite y explote la mente de la imaginacio y creatividad',  
'to both transition careers, and as a hobby',  
'add to my skillset to create personal resilience',  
'To further skills in more advanced design such as algorithms and data structures',  
'I am a college student from Bangladesh.I want to be a computer scientist .So i am learning coding.',  
'I Love Coding and its part of my day activity',  
'I already code. Just wanted to check out this site. Great for beginners but I dont even watch each video and submit the correct each time. Not good for us advanced but if your a beginner then for sure this is a wonderful site for you.',  
'intellect',  
'To see if I enjoy it and want to pursue this as a career.',  
'To revolutionize healthcare',  
'To build new skills to get a job I want',  
'to discovery a new world (curiosity and maybe that it will help me later)',  
'Technological innovation', 'To get a job',  
"Not really sure what I want in life career wise or what interests me but I have seen a few you tubers who code for a living and I've found it interesting and one YouTube recommended this as a good beginning place for coding ",  
'As a hobby that helps me advance my career and to create art/entertainment to share with the world',

"to fulfill God's purpose for my Life",  
 'to develop any software/applications which will be useful for the society',  
 'I love to learn multiple skills so I am constantly prepared for whatever is thrown at me.',  
 'My man money would flow',  
 'To complete my computer skills (i already have a BS in a very similar field) + start a career with it',  
 'PODER CAMBIAR DE DEPENDENCIA EN MI TRABAJO',  
 'To get back into coding',  
 'to help my future once i am older because i am still 12 years old',  
 'Expand my learning of another language.',  
 "To make a Friday Night Funkin' mod :'3",  
 'To add knowledge to a future engineering career. (I am in 10th grade)',  
 'Because it is fun ', 'To add a new skill',  
 'both answers apply - to succeed in current career and as a hobby',  
 'I like coding and want to make a career out of it',  
 'To create my own website',  
 'To modify my career, I am a chemist and want to get into computational chemistry (learning coding and molecular modeling) and chemistry education (learning coding, animation, learning strategies, etc.)',  
 'To continue on my path to learn',  
 'To gain a valuable skill I might have some fun with and make myself more attractive on the job market',  
 'Siempre quise y no me animé, ahora es el momento para mí, y estoy feliz de introducirme en este apasionante mundo',  
 'To integrate my interest in developing tools with my science (oceanography) degree',  
 'To see if coding is right for me',  
 'To be able to contribute to Open Source projects ',  
 'to see if it is for me', 'and also a hobby:)',  
 "All of those things. Basically I've come to a point in my life where I know a bit about computers from my personal hobbies, like playing online computer games, and what I do at work. I have a background in computer network engineering and I sort of just feel like I need to do this in order to succeed at the level that I want. I know it will open doors for me to express my creative side and hopefully do something good for the world as well. ",  
 "It's a mix of reasons. Creating art/ entertainment was my initial reason. I guess you can also say for business and career reasons; because this is something I want to do for life and build on as well.",  
 'I am currently a CS student but I want to supplement my studies ',  
 "as an asset to my career, I'm a curriculum designer.",  
 'Para tener herramientas de desarrollo que me permitan llevar adelante ideas y proyectos, desde la matemática. Y por otro lado que me brinde una salida laboral alternativa como independiente.',  
 'To be useful and versatile inside a company.',  
 'I want to become a programmer developer and more...',  
 'For a better future',  
 'To understand the time I live in and the one to come.',  
 'It is one of my goals in life',

'Durch meine Krankheit CFS (chronique fatigue Syndrom) muss ich etwas neues lernen, da ich leider körperlich meinen Handwerksberuf nicht mehr ausüben kann. Am PC bin ich aber recht fit, so dass ich mir zutraue, programmieren zu lernen.',  
'To become more marketable and also become more well rounded',  
'to change the world',  
'I would like to make money on this and in the future go to another direction in programming',  
'to program activist websites (climate change)',  
'to build quizzes, websites and landing pages and lead magnets for my business',  
'To start a career and to enjoy being a coder',  
'Growth, wanting to someday achieve a specific project',  
'First career/Start business or freelance/hobby/create entertainment',  
'change careers, start a business, hobby, and create video games',  
'Challenge myself', 'Scientific curiosity',  
'To change careers but also to build a portfolio website as a hobby (animation portfolio website) ',  
"just for fun, because i'm a kid",  
"It's not one answer to start my career, as a hobby, and to be able to start my business in the future.",  
'Creating a real god',  
'Bored at home and seeing if it is something I might like',  
'Not only is it a hobby, but to me, coding is extremely fun to watch the time go by.',  
'Plan B', 'To Build my Own Blockchain Network',  
'Because it is the best on the planet!',  
'To learn a brand new skill, and add it to my personal and business resumes.',  
'To meet my dream career requirements', 'To learn new things',  
'to meet expenses for my studies myself',  
'To innovate something or build a product for current world need.',  
'First of all as a hobby and I want to build a website for my relative so that they could teach online.',  
'I am substitute teaching in a middle school coding class, and need to learn how to code to help the students.',  
'oddly, to save \$80/year by diy-ing a popup on my site rather than paying somebody else',  
'understanding life ', 'Getting a jumpstart before college',  
"worried about my future job opportunities, to learn something 'useful'",  
"I want to make a study programs through technology for unlucky kids in Africa, after my college(I'm still in my 11th year high school)",  
'To start my first career as well to change careers from engineering',  
'Hobby , love and money...', 'To expand my skills',  
'For innovation in the tech and IT industry',  
'To widen my horizons/learn a new skill',  
'To change careers and / or as a hobby / or and to start freelance',  
'To upskill myself and for a possible career in coding',  
'To do something for society',  
'I just want to learn it for myself.',  
'next possible step after retiring.',  
"I think I'd enjoy coding and to start my first career",

'I am 8 years from retirement, I want to be able to be relevant after I retire. I have a background in IT infrastructure support and would like to still be able to do stuff IT related in a more relaxed way.',

'To Restart my career', 'To apply inside my future plans.',

'Seeing if I want to change careers',

'recognise its usefulness as a skill to be acquired',

'To expand my job options', 'School & future requirements.',

'curiosity of how thing works, ability to change/create things and solve issues, started as a hobby and became the passion',

'To gain knowledge',

'To get a succesfull career and move to a big city.',

'To prepare for my future career', 'Out of curiosity',

"To help poor people who don't have food.", 'financial freedom',

'At first it was a hobby, now hobby and first career',

'Intento reinventarme y aprender nuevas cosas',

'My reason was to get a job/start a freelancing without a college degree..',

'To get involved with the web3 revolution. To build a better web',

'Skill acquisition ', 'To begin a new path in life.',

"to start a career, but also because it's so much fun and was never exposed to it before ",

'To better understand cryptocurrencies ',

"Feel like it's becoming a necessary skill",

'to meet internship requirements', 'Cualificación profesional.',

'to keep updated',

'to have something to go back to if i lose a job',

'To teach my students. ', 'to add value in my current career',

'to stay updated',

'To change careers, start my first career and/or start a business/freelance.',

'Because I enjoy solving problems and getting paid for it would be nice',

'It always facinated my and I want to try it out before i procced to further education',

'Its important for the future', 'To start programming robots ',

'To succeed either in my current career or to change my career',

"Because it's interesting", 'Passion',

'To solve daily challenges and development',

'This is what I want to study in school so I want to have an edge over the others and also be good & succesful in what I do',

'I want to do it for my own personal achievement ',

'Always wanted to try coding', 'Pure knowledge ',

'to have an alternative career or alternative or additional means of income',

'To make myself more appealing when applying to college',

'and add it to my graphic design portfolio',

'Planning for retirement',

'Learn a new skill to boost current career',

'to be relevant in the future and be able to solve problems with computers',

'To help the Bitcoin Network grow',

'I love to play with computers',

'So many reasons, like those choices above. ;) ',

'looking for my direction', 'too be more versitile',

'To improve techniques and get better jobs',  
'To develop skills for the future',  
'Starting as a hobby and hopefully will become a career one day!',  
'to develop myself',  
"i can't explain my feelingd towards coding..may be i am not good at it..but  
still i love doing it..",  
'Mish-mash of reasons - bored/change careers  
(maybe)/freelance/entertainment/etc.',  
'to improve my income', 'To achieve an intellectual goal',  
'Multiple: To start my first career, to start a business or to freelance, to  
create art or entertainmnet and also meet the school requirements',  
'To learn new things ', 'Dropped out of university',  
'To put it in resume', 'Develop a web application',  
'Both to start my first career and to create art and entertainment',  
'address my ignorance(s) ',  
'Personal development and satisfaction plus potential opportunities',  
'New projects/responsibilities may make it a necessity ',  
'to fulfil a dream',  
"Maybe to start a business or freelance. But immediately, to design a mobile  
site for a friend's business",  
'change career + start 2nd business / Game Dev',  
"It's a hobby, I'm a computer engineering student and I really like it since it  
develops my thinking.",  
'TO LEARN A NEW SKILL', 'fun', 'to move out of my country',  
'To change World', 'To test the water and prepare for university',  
"it's a means to several ends. I want to have more options when it comes to job.  
My current job as a manual software tester is going the way of the dodo, and coding  
could help me in my hobby of game design.",  
'To invent',  
'To start my first career, to freelance and create MMRPG games.',  
'to absorb further knowledge',  
'improve my knowledge about software development',  
'To experience coding',  
'hobby, first career and getting more knowledge to new things',  
'To add up to my skills', 'I like it',  
'To build something useful. To solve crucial problems. And of course to earn  
money. ',  
'To change the world in a good way.',  
'Carrer, Hobby, and To Create Art/Entertainment',  
'I am interested in Coding ', 'I have be interested in coding',  
'retired lifelong software developer wanting to learn the things I missed',  
"First and foremost as a hobby (just something I'm interested in), but I do know  
it might help me in my current or future jobs (e.g. starting my own business).",  
'To meet school requirements, study for school, learn the basics of code to  
help with starting projects and to study ahead of time ',  
'It is something I love to do. It bring happiness whenever I get into coding. ',  
'To future-proof myself and stop being redundant at work (replaceable)',  
'All except for the last one',  
'As a hobby and to have a new skill',



'Coding is a mandatory skill which gives us leverage',  
"To develop revolutionary tech solutions of world's problems",  
'I just love computers and knowing the ins and outs of things ',  
'I enjoy ', 'Struggling to get a job. ', 'To Challenge myself',  
"I would say to make my career easier but it's not required to be successful  
(yet). I do enjoy it as a hobby though.",  
'To change careers, make more money and not be chained to an office',  
'Hobby/find career',  
'Because all of my music fan pages were hacked',  
'As a hobby and possibly career applications',  
'Change careers. I used to be an Aviation Technician, Office Manager. I am now a  
caregiver to my disabled son. I want a different career now.',  
"I have always wanted to learn coding but I didn't know where to start. That's  
when I heard about freecodeacademy, they helped me get started with it! and I'm already  
loving it.",  
'learn coding before college',  
'because i love coding and i want to spend all my life coding and learning new  
things',  
'to learn how to make a decent website',  
'I am not sure yet - personal project but it would be great if it ends with a  
nice job?',  
'change career, start career and build skill',  
"To give my family a better life and I also want to create a better future for  
me, I'm from latin america and here there are not many opportunities to improve our  
lives.",  
'Construction work has been rough on my body.',  
'to have the skill at hand if i have a great idea for an app, website or piece  
of technology.',  
'Want to create an OS like Winddows or MacOS',  
'I want to learn the technology and know how big companies work and make that  
knowlege available to all. ',  
'As a hobby, to meet school requirements, to start a business or to freelance,  
and to start my first career (I have a lot of reasons to learn to code)',  
'To broaden my skillset',  
'Hobby that I hope to turn into a first career through business/freelancing. ',  
'As a hobby, for school, and maybe to start a career',  
"I'm 16 and I'm trying to figure out what to do in the future and the first  
thing i thought of was being a web designer.",  
'for meeting school requirements and more interested in learning more',  
'Used to program in c years ago. got lost is visual programming,so quit.',  
'the answer to this question would probably be too long and complicated for this  
field',  
'To venture into a new domain learning',  
'I want to start my career as a freelancer and to run a business',  
'to have a complimentary discipline',  
'Broaden my skills/knowledge',  
'Curious to know about the workings of computers; and interested in code +  
career oriented plans.',  
'I cant find an app to meet my needs, so I want to make it for myself.',

'To invest in the future.',  
'I love technology and i want to make money from it',  
'To restart my programming career after a long break',  
'Porque me gusta ', 'be a more informed person',  
'To broaden my skillset and expand available career options',  
'expand knowledge for current career',  
'As a hobby, possibly for a career change',  
'To expand current career',  
'fascination and career/freelance potentials', 'knowledge',  
'I want to learn to code.',  
'as a hobby and to start my first career',  
"It's been like a secondary occupation for several years",  
"I'm a stay at home mom with lots of time on my hands.",  
'Because I really want to be great at It',  
'Change careers, coding makes me happy, to build a dating app for my sister. ',  
'I want to make cool stuff that solves problems people have. For my current job and/or next job.',  
'To teach CS to high school students',  
'To keep myself learning and busy', 'interest',  
'self fulfillment i guess its just something that really intrigues me',  
'To help solve problems', 'To improve my skillset', 'For fun',  
'get back into workplace after raising kids',  
"I first had the desire to learn how to code way back in 1996 at Choptican High School. I signed up for Computer Science and got KEYBOARDING and office skills. I am an African American female with three children I am 40 years old I've long had a passion for technology and often found myself ushered in a different direction. President Obama had a day of code in 2012 and I sat down and learned the concepts of HTML on that day I decided I wanted to learn how to program so that I could help women like myself work from home and be able to support their children and help their community. Everybody doesn't want to be a nurse everybody doesn't want to be a secretary at the same time a high-dollar professional skill is the only thing that can help a woman rise and shatter glass ceilings and take care of her family in this economy. I want to learn HTML CSS and JavaScript back in a python mphp so that I can BUILD my Shopify e-commerce store I already have markups and products for my store I want to learn the skills to start my own business I don't want to have to pay somebody to do something that I've full well can learn myself in desire to. With that high-dollar skill that will help me open my own business on the internet entering the world of e-commerce no later than 2022 I also plan on freelancing that highly in demand skill. My 15 year old son and I are both learning python HTML CSS and JavaScript on Coursera we also use w3schools as well as sololearn we are learning Shopify developer skills through codingphase as well as coding with mosh we learn various different other languages I used all of these teachables and free apps to have a curriculum that well inculcate the principles that I'm learning I really want to learn computer science but I don't have the money for college and I don't have the time for college right now but I need to learn so then utilizing the skill that is necessary in programming self-motivation to learn the skill of program, I want to be grounded in the principles. The ultimate goal is to start an e-commerce SEO Shopify Development firm. There's so much marketed to black and brown children on YouTube and the internet, and I don't see many coding teachables for BIPOC children, poor internet service is teaching young by pop children to code that is even

within their reach all those kid coding camps prices are so astronomical and insane you can totally see how the wealth disparity will continue! So my son and I plan on making our own teachables geared toward children of color and broken down in a way that they can understand, grasp, and put into action with the resource that we are learning on Python for Everybody on Coursera. I want to break the generational wealth curse, through my family's own volition, sweat, hard work, and action... with the help of Jesus and many smart people kind enough to impart their wisdom, as well. Lol.",

'to get a stable job and income, once financially aided, planning to open a business/start investing.',

'to explore the world of technology everyday something new and sophisticated',

"I'm a professional software engineer who has gotten rusty while transitioning into a management role",

'Equally to start first career and as a hobby',

'I feel left out when people I admire want a website or app team',

'I keep me busy because I start doing stupid things',

'I have always had an interest in computers so it's sort of natural that I drifted to coding ',

'career and hobby',

'To run away from dirty and cold job outside in winter. To become smart and rich. To always stimulate my brain by learning more.',

'Side hustle',

'to create things that life exciting. hopefully friendly, intelligent robots.',

'To be a useful human being maybe',

'To have more options in my current career path. Also prepping for an MBA next year and consider this will be a great add on!'

"Because my sociology degree didn't work out. ",

"To continue what I've started and be flexible with my time and to continue the abandoned passion and also to be effective in Jesus' ministry!",

'First, to do better in my job. Second, to have a backup plan in case my job falls apart.',

'To see if it is right for me',

'Because it can always be useful and I think that it might really help in my CV even though my studies have nothing to do with computer science',

'Personal Development', 'For many reasons ',

'To contribute to the society',

'To Get Into A Good College Or Software Career',

'I like coding and want to get a job in microsoft. ',

"To see if this is something I'm good at and maybe start a career",

'mixed between creating something of value that maybe someday will change the world, and having fun along the journey .',

"It started out as a hobby but now it's to change career. I am from an Electrical Engineering background and now I want to switch to computer science background.",

'To master in what i have learnt..',

'To learn with my kids who are 4 and 8 years old.', 'Conocimiento',

"Always wanted to created my own product, and to solve people's problems in other industries.",

'because it very enjoyable and to start new career',

'To teach my son',

'I wanted to experience coding first to see if I was up to the task and could do this before I make a real commitment and so far I like what I have done and experienced so far.',

'As a hobby and to be diverse in current career',

'to make myself more useful and to prep for my future career',

'To make program files, then to make useful ones, then to get a career in it',

'Game Development', 'video games and a hobby',

'makes me happy and i feel like god',

'To build a retirement income',

'Will be retiring in a few years and want to continue working, from home this time.',

"I've been addicted to code since Basic!",

'Retire and begin a second career.', 'to succeed in school',

'Software development',

'To offer better solutions to my customers',

'To learn more about computer science.',

"It's basically everything above except for change career and start a business. I'm a student and I'm planning to work on a company not start one",

'To Expand my Skills',

'Well, I Just love technologies and in order to achieve my dreams I need to be a coding master ',

'a little bit of both, starting my first career and to create art/entertainment',

'to do some random useful cool stuff whenever possible',

'Broaden knowledge and improve earnings', 'To Innovate',

'Mainly I want to start my own website!',

'Both as a hobby and do eventually make a career change.',

'curiosity', 'To create an app to help people with disabilities',

'To convert PureData abstractions to C++, de-mystify code and change career',

'To expand my learning base for the upcoming world',

"i'm still in high school, but i wanted to start studying to code early, so in university i'll know already something",

'Learning new skills',

'one day i would like to leave corporate world for non profit world',

'To build things to help others',

'To enter a new career and to acquire skills for video game development',

'to find something that motivates me and to change careers',

"It's my calling",

'All of the above, except meeting school requirements',

'To carry on with my career',

"to see if coding is something I'm interested in",

"It's the first time I'm finding myself staying consistent despite numerous challenges",

'for fun as a hobby, but also to create something other people can enjoy and use',

'Personal Growth', 'Further my technical skills',

"I work 60 hours a week and still don't make enough to get ahead.",

'It's my passion and I love it ',

'to get necessary skills for a passive income',

'Already know how to code, looking to contribute more to Github and want to

understand how others are learning.',

'Due to my country(Turkey) rapidly devaluating currency. I am seeking jobs that can offer me payment in US dollars/Euro currency.',

'to change my career, because this year i was always rejected trying to enter this career. Aswell as a hobby, because i want to start modding or even creating my own games, because i have many ideas and so little to realise my projects at the moment. It would be more than doing my own game, it would really affect my life in a way i need.',

'As a qualification/resume builder',

'to understand world in digital era',

'To see what it's like. Pure curiosity.', 'For a school project',

'Cambiar mi vida y la de mucha gente',

"to relearn things I've forgot and change careers",

"I took it to prepare for a coding class that is required for my marketing degree. I soon fell in love with coding. Now I'm going to change my career path to web development!",

'To apply change in the world.',

'To start my first career and as a hobby',

'To build personal projects',

'to support my studies in computer science',

'To build products myself instead of only designing them.',

'New skill',

'To understand the Cryptocurrency movement on a deeper level.',

'Hobby, change careers, simply wish to learn',

'Requirements, business and creator',

'To be successful in my life',

'To work on a project that needs these skills',

'A backup plan for my current career',

'To see if good for people I mentor',

'Because I love tech and creating things, make my imagination real. Programming is an art from which you can develop something which can be used by billions of people for their daily life enhancement and optimization.',

'I love to learn new things and this could help me get a job later in life. I also love computers, I take them apart and put them back together. Because of that I wanted to learn the language of computers. ',

'enrichment on the side but hoping for career success too',

'To adapt to higher computer dependency and usage',

'To change the World! :)', 'Plus for money too',

'Passion to learn coding',

'To see if it something I want as a career',

'To learn how to program the robots in my warehouse instead of get my job taken over by them',

'Just to learn', 'Practice old skills and develop new ones',

'Both the second and last question '], dtype=object)

As you can see, the respondents gave a lot of different answers to this question. To get a better idea of the most common responses, we will plot a graph which tells us the most common reasons for learning how to code.

```
biggest_reasons = (survey_df.BiggestReasonToCode.value_counts(normalize=True)*100).head  
biggest_reasons
```

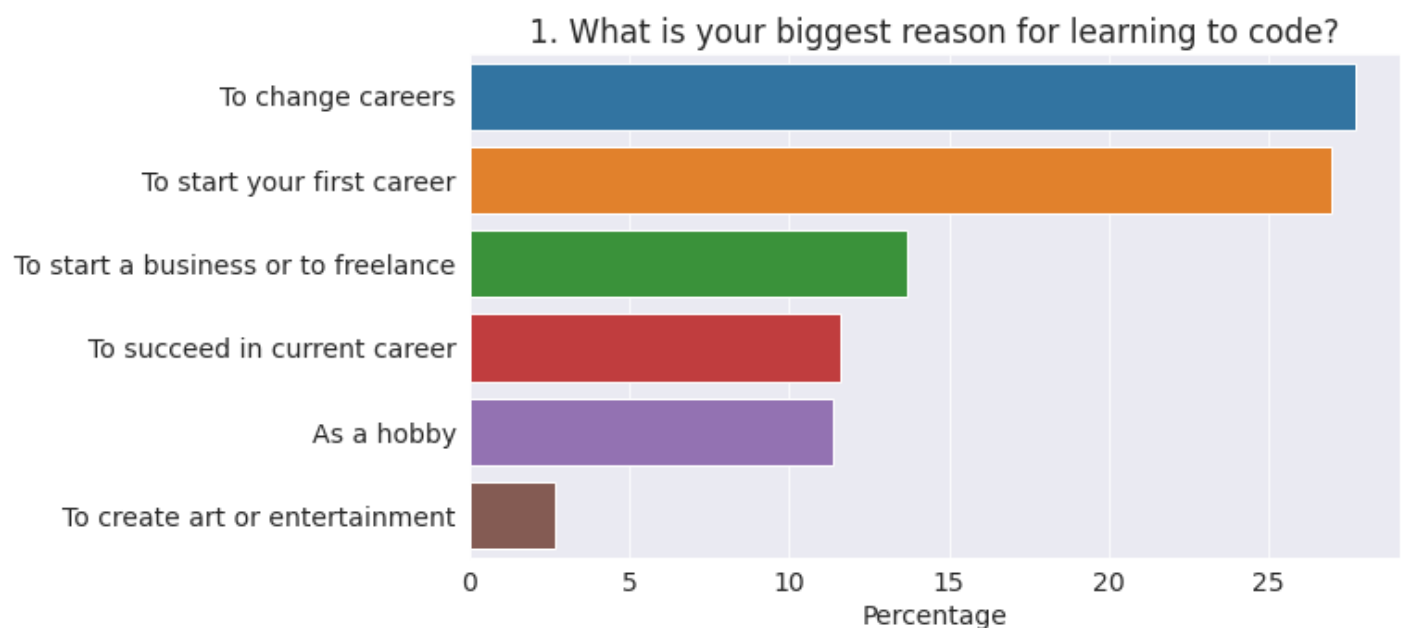
To change careers	27.696023
To start your first career	26.949000
To start a business or to freelance	13.699169
To succeed in current career	11.592900
As a hobby	11.385082
To create art or entertainment	2.684790
To meet school requirements	1.561447
Curiosity	0.022467
To change the world	0.022467
All the above	0.011233

Name: BiggestReasonToCode, dtype: float64

```
sns.barplot(biggest_reasons[0:6], biggest_reasons.index[0:6])
plt.title(questions['BiggestReasonToCode'])
plt.ylabel(None)
plt.xlabel('Percentage');
```

/opt/conda/lib/python3.9/site-packages/seaborn/\_decorators.py:36: FutureWarning: Pass the following variables as keyword args: x, y. From version 0.12, the only valid positional argument will be `data`, and passing other arguments without an explicit keyword will result in an error or misinterpretation.

```
warnings.warn(
```



We can see clearly from the results the 2 primary reasons people begin learning how to code is to either change career path they are currently on, or to start their first career, which makes a lot of sense as a lot of people chosen to do this survey would be new or aspiring coders/developers.

A lot of the most common responses are focused on career building, but there are also a lot of people (roughly 12 %) who choose to learn to code solely as a hobby.

## Gender of respondents

Data analysis/Data science are considered very male dominated fields. We can take a closer look at our data to validate whether or not this statement holds up.

```
questions['Gender']
```

```
'24. Which of the following best represents how you think of yourself?'
```

```
survey_df.Gender.unique()
```

```
array(['Prefer not to say', 'Male', nan, 'Female', 'Nonbinary',  
      'Nonbinary woman', 'Queer', 'Gay',  
      "Physically female, but I don't relate to that (or any gender) as an identity",  
      'trans', 'Trans male', '18', 'Traditional Male', 'Trans guy',  
      'Male, He/They/Them', 'Living being ', 'I am a meat popsicle',  
      'I see myself as a person who is not restricted not limited by my gender as a  
female in being the best that I can be per time being blessed and being a blessing  
(problem solver/solution provider)',  
      'Transmale', 'Transmasculine Genderqueer', 'Nonbinary girl',  
      'FTM Trans',  
      'Assigned male at birth. Female. I am a trans woman.', 'Human',  
      'genderfluid or demigirl', 'Diverse and flexible', 'He/Him',  
      'Male teen', 'PINEAPPLE (rpz le meilleur sex)', 'Agender',  
      'transgender', 'Mixed', 'Gamer', 'trans man', 'Gaki jk (male) ',  
      'Female High-schooler', 'Trans female', 'Trans Female',  
      'not supposed to ask personal info', 'trans male',  
      'a human person :) (male)', "this question doesn't matter",  
      'A guy who is determined to learn web development',  
      'Attack Chopper', 'Alpha male', 'person',  
      'Learn always new subjects', 'non-binary woman', 'student?',  
      'Male, just Male', 'questioning ',  
      'possibly a descendant of a Norse God.', 'Ajith', 'Just me',  
      "I'm human. What else matters?", 'bi sexual', 'transguy',  
      'Trans Girl', 'Carbon-15', 'Life', 'attack helicopter', 'universe',  
      'I am a woman but "female" and "male" are not the best terms - you should use  
man or woman since you\'re asking about gender not biological sex. Also, the term  
"female" is often used in a dehumanizing manner. ',  
      'pre anything MtF',  
      'I do not belive in people emotions and desires that oppose nature(i am male)',  
      'Motiviert, zielstrebend und zuverlässig', 'demigirl',  
      "What the actual fuck. I'm male.", 'There are just two genders',  
      'WTF', 'Demiboy', 'Transgender man',  
      'Male but I like how you phrased that question :P',  
      '(Trans) Woman', 'Trans-NB', 'fluid', 'Attack Helicopter',  
      'Alpha ;)', 'Straight forward lol.', "j'ai compris la question",  
      'Genderfluid',  
      'I don\'t "think of myself" as anything, I am a male.', 'alien',  
      "Transgenderism is an abomination, and is a part of the repeat of the Nazi's  
rise to power. G-d made only men and women, there is nothing else and it is a  
perversion and a delusion to say otherwise. This movement is really about child  
molestation and mass murder. Your organization is wicked for promoting the agenda of  
Nazism. The Nazi party were child molesting homosexual murderers and they still are.",  
      'He/They trans FTM',
```

```
"LOL how dumb. Why do people get sucked into this nonsense? I'm giraffe okay!?",
'A Creative, compassionate citizen of the Global Garden.', 'Spock',
'Strong', 'bi', 'Bigender She/her he/him',
'Genetically and scientifically male', 'Am a human not an alien',
'Trans Masc ', 'Michelle "Big Mike" Obama', 'Homosexual'],
dtype=object)
```

```
gender_counts = survey_df.Gender.value_counts()
gender_counts
```

```
Male                12074
Female              4456
Prefer not to say    327
Nonbinary           320
Genderfluid          4
...
Gamer                1
Mixed                1
transgender          1
PINEAPPLE (rpz le meilleur sex)  1
Homosexual           1
Name: Gender, Length: 95, dtype: int64
```

There are a lot of responses which occur fewer than 5 times. We will focus on the 3 most common responses (not including those who chose not to say), which are Male, Female and Nonbinary.

```
most_common_genders_df = survey_df[survey_df['Gender'].str.contains('Male') | survey_df
```

```
most_common_genders = (most_common_genders_df.Gender.value_counts(normalize=True).head(
most_common_genders
```

```
Male                71.613286
Female              26.429419
Nonbinary           1.897983
Name: Gender, dtype: float64
```

As you can see, we have created a new df with entries that only contain 'Male','Female', or 'Nonbinary'. It's not perfect as there is some overlap with the responses, but it still gives us a good idea of the distribution. We can now plot this.

```
colors = sns.color_palette('pastel')
plt.pie(most_common_genders,
        labels=most_common_genders.index,
        colors=colors,
        autopct='%.0f%%')
plt.title(questions['Gender']);
```





```

round(100-sum(common_ethnicities),2)]

common_ethnicities_list = [common_ethnicities.index[0] + ' ' + '(' + str(round(common_e
common_ethnicities.index[1] + ' ' + '(' + str(round(common_e
common_ethnicities.index[2] + ' ' + '(' + str(round(common_e
common_ethnicities.index[3] + ' ' + '(' + str(round(common_e
common_ethnicities.index[4] + ' ' + '(' + str(round(common_e
        'Other' + ' ' + '(' + str(round(100-sum(common_ethnicit

print(common_ethnicities_list)
print(common_ethnicities_pct)

```

```

['White (36.23%)', 'South Asian (14.31%)', 'Hispanic/Latino(a/x) (14.26%)',
'Black/African American (11.4%)', 'Southeast Asian (9.44%)', 'Other (14.36%)']
[36.23, 14.31, 14.26, 11.4, 9.44, 14.36]

```

```

pip install squarify # we need to install this to create the treemap

```

Collecting squarify

Downloading squarify-0.4.3-py3-none-any.whl (4.3 kB)

Installing collected packages: squarify

Successfully installed squarify-0.4.3

Note: you may need to restart the kernel to use updated packages.

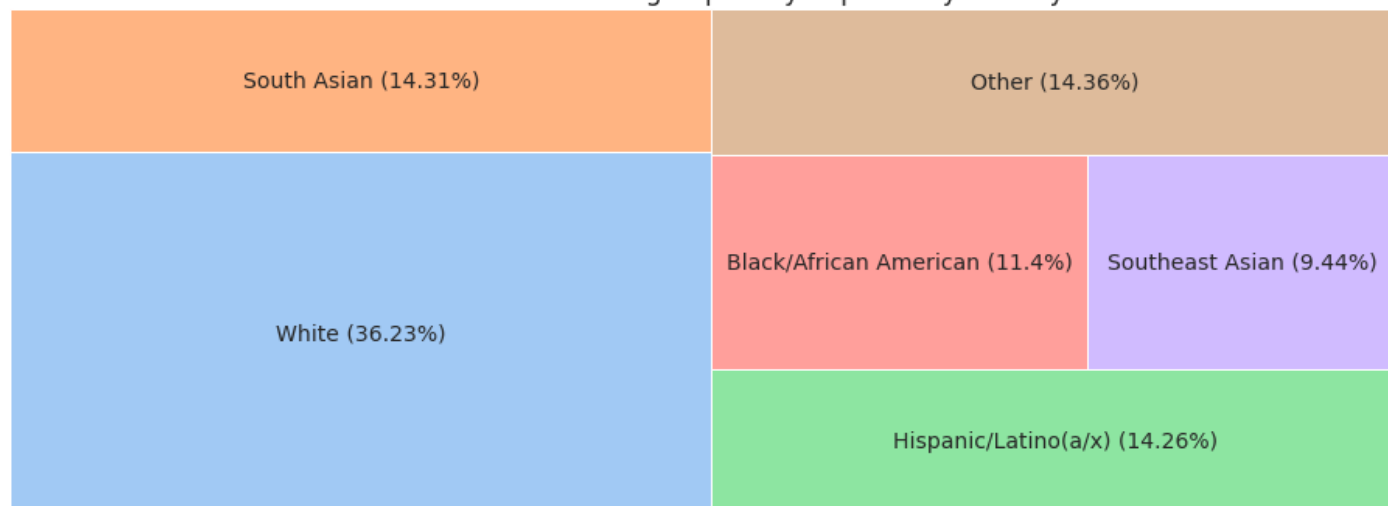
Now we can create our treemap graph.

```

import squarify
colors = sns.color_palette('pastel')
plt.figure(figsize=(16,6))
squarify.plot(sizes=common_ethnicities_pct,
              label=common_ethnicities_list,
              color=colors)
plt.title(questions['Ethnicity'])
plt.axis('off');

```

25. With which of these groups do you primarily identify?



More than a third of the respondents were white, so whilst there may be a slight bias towards predominantly white nations, there is still a good amount of representation from other backgrounds in the dataset.

## What the respondents studied

```
questions['UniversityDegree']
```

```
'33. If you attended a university, what did you study?'
```

```
survey_df.UniversityDegree.unique()
```

```
array(['Computer science, computer engineering, or software engineering',  
      'Environmental science (e.g., earth sciences, sustainability)',  
      'Art',  
      'Another engineering discipline (e.g., civil, electrical, mechanical)',  
      'A business discipline (e.g., accounting, finance, marketing)',  
      "I didn't attend a university", nan,  
      'Information systems, information technology, or system administration',  
      'Fine arts or performing arts (e.g., graphic design, music, studio, art)',  
      'A natural science (e.g., biology, chemistry, physics)',  
      'A health science (e.g., nursing, pharmacy, radiology)',  
      'Data Science', 'Education, then psychology', 'Asian studies',  
      'A social science (e.g., sociology, psychology, political science, economics)',  
      'Am currently joining the university to study Bsc.Software Engineering',  
      'Event Management',  
      'A humanities discipline (e.g., literature, history, philosophy)',  
      'Mathematics or statistics', 'Undecided or no major',  
      'linguistics', 'Web development or web design',  
      'about to start SE, if my wallet can bank it ... got some few college credits  
unrelated to the degree',  
      'Will attend soon', 'Paralegal', 'Education ',  
      'BA in biological sciences, 2yr post grad study in environmental sciences, MA in  
clinical psychology',  
      'Architecture', 'Law', "2 bachelor's: business and IT",  
      'Kinesiology ', 'I am yet to attend a university', 'yet to attend',  
      'Electrical engineering ',  
      'Applying this fall, looking in majoring comp sci',  
      'Architectural Technology', 'Health Service Administration',  
      'Education', 'law', 'Computer science & health&safety science',  
      'Agriculture', 'Multimedia',  
      'First business administration, then CS; quit both halfway through',  
      'Engineering Physics ', "I'm still in highschool", 'Linguistics',  
      'Still studying', 'Currently in university', 'Translation',  
      'Yeah liberal arts...lots of stuff from the above list.',  
      'Communications (e.g., journalism, TV production)',  
      "I'm currently studying a Computer Science BSc degree",  
      'Languages', 'Ancient Greek', 'Geotechnical engineering',  
      'Spanish/CompSci',  
      'planning to attend after high school and major in computer science',
```

'architecture', 'Education - Teaching K-12',  
 'Civil Engineering, Physics and Statistics', 'Laws', 'anthro/econ',  
 'Visual Communication',  
 'Kinesiology (Exercise Science & Health Promotion)',  
 'Environmental science technology and management year one',  
 'Digital Media and Design',  
 'Interdisciplinary (English Communications & Environmental Studies)',  
 'Still in secondary school', 'Computer Science-Economics',  
 'Physics', 'Technology Engineering & Innovation Management',  
 'nutrition and food science', 'tourism',  
 'Currently at the last year of High School',  
 'Humanities BA, and then MDes', 'General Studies',  
 'Legal Studies ', 'Chemical engineering',  
 'currently doing a degree in information technology',  
 'criminal justice', "Haven't attended but after college I will.",  
 'Sports Management', 'BA in social science, MPA in business',  
 'Agricultural Extension And Communication Technology ',  
 'Accounting', 'IT ', 'electrical engineering',  
 'Roughly translated it was called Media Computer Science and Design. I also  
 studied Sociolgy and Philosophy.',  
 'Bachelors of Computer Applications', 'Industrial design',  
 'Still in high school', "I'm still in elementary",  
 'teaching degree', 'Criminal Justice', 'Mechatronic',  
 'Quantity surveying', 'cyber security',  
 "I've switched from Anthropology to neuroscience to (currently) Computer  
 Science",  
 'Physical Sciences ', 'commerce', 'gen ed',  
 'about to attend university for cs', 'game design', 'theology',  
 'English lit and lang', 'Electronic engineering', 'Social work',  
 'media',  
 'Planning on going to uni in 1-2 years thinking about studying computer  
 science',  
 'Civil Engineer ', 'Modern Languages', 'Biology',  
 'I am yet 15, still in high school',  
 'currently in University studying Biotechnology',  
 'Computer and Information Technology major', 'Management and Law ',  
 'Business information technology',  
 'I want to study Chemical Engineering',  
 'mathematics and financial engineering',  
 'I am currently studying High School',  
 'i am currently studying computer science', 'CSc but dropped out',  
 'Not yet', 'Still Awaiting Admission ', 'I wont attend one',  
 'Criminology and Police Science', 'b.tech(Mechanical engineering)',  
 'Military Leadership', 'Homeland Security/Emergency Management',  
 'In my vocational institution I study Arts and Graphic Design, but it is not  
 university',  
 'operations research and statistics', 'BS biology + minor math',  
 'I am currently studying a university degree in national gendarmerie on citizen  
 security with guidance in telecommunications.',

'city and regional planning or urban planning (was within the engineering school, but other university programs are within their design school)',  
 'Administration ', 'law enforcement',  
 'Islamic Theology and Arabic', 'finishing high school now',  
 'Computer science, computer engineering, or software engineering, engineering for robotics',  
 'Three degrees: Business, Humanities and Education',  
 'I am still in highschool. About to be a senior',  
 'Teaching of German',  
 'Classical philology, Indoeuropean linguistics, Comparative Religions (was doing a mixed 5-year degree)',  
 'Elementary Education',  
 'Software Development and System Administration',  
 'I am in high school.',  
 'tecnicatura en programcion (desarrollo de software)',  
 'Not yet in university',  
 'Am attending college for Heavy Duty Repair', 'Not old enough yet',  
 'legal', 'Undecided/Fine Art, looking to switch to CS',  
 'Mortuary Science', 'Not in uni yet',  
 'Hotel and Restaurant Management', 'Computer technology',  
 'Cybersecurity', 'Childhood studies/education', 'Electronics',  
 "I'm attending the university now in computer science",  
 'International Hospitality', 'library management', 'Biotechnology',  
 'Student ', 'i am still in highschool!', 'Logistic',  
 'Air Traffic Control',  
 'criminal justice (not considered a social science degree)',  
 'Science Education', 'Bruh not old enough for that',  
 'Business, then a certification in web development. ',  
 'currently attending and majoring in computer science',  
 'Computer Applications', 'Art Education', 'Public Administration',  
 'Kinesiology', '(Socio)linguistics and Chinese',  
 'International Business',  
 'sociology b.a. with a media studies minor',  
 'Mechanical Engineering', 'ingenieria electronica ',  
 'Double major in Neuroscience and Cognitive Psychology',  
 'Physics and Math mostly',  
 'Waiting for university acceptance results',  
 'Have degrees or minors in Business History Political Science and Comparative Culture studies ',  
 'English Language ', 'Business and Commerce',  
 'universities are useless and outdated when it comes to practical degrees like engineering,MBA ,etc which can be learned only through real world jobs. for science or math or social science its a different story, for that u need higher degrees(like MS/Phd)but for practical subjects its better to have experience and skill before taking masters(if required [too costly] ).',  
 'A social science (politics) and computer science',  
 'Still studying Bachelor of science in Quantitative Economics',  
 'hospitality', 'Hotel Business.', 'Education Faculity',  
 'Traditional Chinese Medicine',

'Audio and recording technology, so somewhere between electronics/physics, and engineering within the musical realm',  
'i am in the uk and i am still attending secondary school year 11',  
'English language teaching', 'Medieninformatik',  
'enviromental sciences and zoology',  
'still in high school want to do computer science',  
'Airline Pilot', 'Civil Engineering', 'chemistry(Edu)',  
'Medical School', 'I am still studying',  
'Artificial Intelligence and Data Science', 'Computing',  
"I'm still in high school, grade11", 'Currently in medical school',  
'Medical Aesthetics',  
'Bioinformatics (1 year), biology (1 year), film school (3 years)',  
'business administration and commerce', "Still haven't ",  
'games design & development', 'Computer Games Art',  
'not yet attended',  
'Business ADMN & Media Studies Joint Major with a Minor in ComSci',  
'Education (General Science and I CT) ',  
'Liberal Studies for Teaching', 'A business of airlines industry',  
'Industrial Engineering', 'german language',  
'Technology education in computer studies', 'Law School',  
'Business Information Technology', 'Not in college',  
'agricultural science', 'Coding',  
'Currently attending a University', 'bachelor of Science',  
'Lawyer', 'b.com (computers)', 'tourism management',  
'Industrial Physics',  
'environmental, safety and occupational health', 'RN',  
'Transportation Management ', 'Liberal Arts', 'Agriculture ',  
'International Relation and Diplomacy',  
'Communications major and business minor', 'medicine',  
'Administration',  
'Liberal arts degree with focus on communications, design and computer science',  
'?',  
'Earth sciences (Have a good mathematics background)',  
'Digital Media / Creative writing', 'Student',  
'Softwarelocalization', 'Economics', 'Network engnieer',  
'Applied science-paralegal ', 'Agricultural Economics with IT',  
' i am still high school student', "I don't attend university",  
'Multimedia Systems Design', 'Agricultural Sciences',  
'Speech and Hearing Science',  
'Peace Studies: Sociology, Philosophy, Political Science, Anthropology',  
'Clinical nutrition and business', 'Business Administration ',  
'Airlines, Tourism and Hospitality Management', 'Urban Planning',  
'International Studies and nursing', 'Human Management ',  
'Attending for CS Major', 'and studio arts courses ',  
'Senior High School', 'Pharmacy',  
'I am curently attending university for sociology',  
'AGRICULTURAL ECONOMICS', 'Sport Science',  
'I am Still Attending an University, Mechanical Engg.',  
'International trade', 'BASIC PRE-REQUISITES',

'Rehabilitation Services', 'Business and Psychology double major',  
'Applied Technology', "I haven't attended a university yet",  
'Human Centred Design', 'Supply Chain Management',  
'Master of Business Administration',  
'just completed my high school looking for collage', 'in college',  
'Music', 'Artificial Intelligence', 'agriculture',  
"haven't attend a university", 'Networking', 'Quantity Surveying',  
'education', 'Secondary Math Education', 'stem',  
'Industrial Technology (manufacturing mgmt)', 'Still in secondary',  
'Religious Studies',  
'software and other engineering, it, web,math or stat, humanities discipline',  
'Entomology', "I'm in the high school yet. ^^", 'still in school',  
'not old enough to attend college', 'Food Science',  
'Biomedical Engineering', 'natural science and business',  
'game dev', 'Software Development', 'Aerospace Engineering',  
'earlier Linguistics, now Human-Computer Interaction',  
'Electrical and Electronic Engineering', 'Hospitality',  
'cyber security science', 'Sociology and computer science',  
'Aviation', 'Industrial Design', 'Chemistry', 'graphics',  
'Triple major: music; physical chemistry; math; and graduated from honors  
program as well',  
'Special Education', 'Computer Science and Pharmacy',  
'meteorology', 'not quite there yet', 'Theology',  
"I'm enrolled in an University for a business and arts interdisciplinary degree.  
The program starts Fall 2021.",  
'I AM IN 7th CLASS CURRENTLY', 'Library and Information Science',  
"Currently pursuing a bachelor's in business and economics",  
'mechanical engineering', 'M.Sc. in Applied Geology',  
'I am attending university this year',  
'mechanical engineering - health science - a natural science',  
'law and management and psychology',  
'Law then I switched to Foreign Languages', 'biochemistry',  
'English', 'Digital Forensics and Computer Science',  
'IT - Game Development', 'biomedical engineering',  
"I'm currently in the university studying Mechatronics Eng",  
'Psychology', 'currently enrolled in secondary school',  
'Lic. Educación Física', 'Physics major, computer science minor',  
'B.Sc Agriculture',  
'Still in highschool but planning on being a biomedical engineer major with  
computer science minor',  
'currently studying finance', 'I am going to but not old enough',  
"I'm a highschool student !!", 'English Literature', 'languages',  
'not yet', 'i am a student in high school',  
'education, Spanish, French', 'medicin', 'Child Psychology',  
'Double Major: Software Engineering and Industrial Engineering',  
'Bachelor of Arts for higher services of the german police',  
'Actuarial Science', 'Many different computer science fields',  
"Haven't attended", 'law school', 'health information management',  
'Speech Pathology, Social Work, Computer Science',

'artificial intelligence',  
 'Social science, humanities, performing arts',  
 'Business +Computer Science Double degree (one by it self narrows your vision)',  
 'Elementary education ', 'Education/ humanities ',  
 'my 3 degrees were 1. computer programming & analysis 2. website design 3. video  
 game developement & when i left school there arent very many jobs for anything i went  
 to school for in this area of the world.',  
 'Physics at the Lebanese University, Faculty of Science',  
 'architecture and urbanism', 'transports, road safety',  
 'Psychology ', "Haven't Attended yet",  
 'Education, Social Science, and Fine Arts',  
 'i am too young to attend', "I'm still in high school",  
 'BCom IIOP', 'I am in the secondary school',  
 'still in high school', 'Analytics', 'Software development',  
 'I was going for Comp Sci and Psych, double major',  
 'Ingenieria de sistemas',  
 'currently pursuing in computer science & engineering major',  
 'Neuroscience and Computer Science', 'i have not', 'Management ',  
 'Mechatronics Engineering', 'Music Technology BSc', 'lawyer',  
 'Culinary management', 'English Language and Culture',  
 'social pedagogy', 'Secondary Education and English double major',  
 'Healthcare', 'Agricultural Economics',  
 "I'm going to start attending University this year",  
 "I'm still studying", 'Education Technology ',  
 'Information systems, information technology, or system administration AND  
 Education',  
 'Geography', "I still haven't finished high school",  
 'Human Resources', 'Language education',  
 'Library and information science',  
 "I'm in the process of application",  
 'Telecoms - Communication Systems and Digital Networks',  
 'A little career of programming.', 'Secondary Education',  
 'Music Business', 'Not yet at a university',  
 'Technical Diploma X-Ray Department ',  
 'International History and Politics, Psychology, Computing & IT',  
 'currently studying a mix of industrial engineering and geoinformation',  
 'I have two degrees - marketing and computer science',  
 "I'm a student studying electronic and computer engineering",  
 'Game Design', 'Accelerators and Detectors Technologies',  
 'I am still enrolled in University pursuing Computer Science & Engineering.',  
 'Art & technology of TV & Film.', 'General Transfer',  
 'I'm still in high school',  
 'Electronics and Communication Engineering',  
 'Electronics and Instrumentation Engineering', 'Civil engineering',  
 'Biology, Mechanical Engineering, Computer Science', 'Language',  
 'Teaching degree',  
 'Computing in undergraduate. Business in Masters',  
 'Physics Education', 'PPE',  
 'I am currently enrolled in a university (mathematics degree)',



'Electronics and Telecommunication Engineering',  
 'Animation and Early Childhood Development', 'Spanish and Math',  
 "I'm just starting college",  
 'architecture and nursing - odd to not be able to choose multiple areas of  
 study?',  
 'Japanese Studies and Japanese language', 'Commerce',  
 'Computer security and forensics',  
 'Marketing, advertising and graphic design ', 'language',  
 'Mechanical engineering', 'information science',  
 'Paramedic School', 'Educational Technology',  
 'Tecnatura universitaria en Programación', ' computer science',  
 'Curriculum design, adult education and training',  
 'Mechatronics Technology', 'Bioinformatics',  
 'Languages and Cultural Mediation',  
 'I will attend soon to computer science',  
 "I haven't attended a university yet.", 'im not in college yet',  
 'CS and Bio Double Degree', 'Business IT', 'Biochemistry', 'LAW',  
 'Computer Science and Electronics',  
 'Financial management Technology',  
 "I haven't attended a university", 'Film & Video Production ',  
 'Food Science and Engineering', 'Horticulture',  
 '1st - Architecture then Master History of Architecture. Now taking History  
 degree',  
 "I am only 15, haven't thought of university yet.",  
 'Currently attending under Pre-Law with Criminal Justice major',  
 'Mathematics and Computer Science',  
 'Currently studying Information Technology', 'I am too young',  
 'physics and computer science', 'Currently attending a university',  
 'Electronic Engineering',  
 'Currently going for business management ',  
 'Biotechnology, a applied natural science',  
 'Bachelor of computer applications', 'i am still in school',  
 'Sports ',  
 'currently doing my MBA, my bachelors is in anthropology',  
 "I haven't attended university", 'liberal arts', 'Digital Design',  
 'Aeronautics', 'Game Developer', 'Book design',  
 'Veterinary medicine', 'Petroleum Engineering',  
 'software engineering', 'Mechanical Engineering ',  
 'Business Administration / Technical Management',  
 'Architecture; Urban Planning', 'Still Studying', 'Games Design',  
 'i dont attend university but plan on doing so after high school',  
 'not joined yet ', 'languages, Finance, Accounting',  
 'I have studied Electrical Engineering and a Natural Science and am currently  
 enrolled in Computer Science',  
 'Agronomy', 'Classical languages/linguistics',  
 'took communications before dropping out now taking information tech',  
 'I will attend university from next year', 'finance and banking ',  
 'physics', 'Independent major - music education for south Asia',  
 ', ', 'Criminal justice', 'science', 'Mining engineering',

'Art and Design', 'will be attending university next year',  
 'applied science', 'Business Management', '3D animation',  
 'Recording Arts & Music Business',  
 'Higher technical degree in programming, incomplete',  
 'Digital Design Production', 'Business Management ',  
 'Mechanical engineering', 'Medicine',  
 'Business Intelligence and business analytics', 'Microbiology',  
 'Management', 'My university starts in 3 days',  
 'Electrical engineering + Communication + Web dev./Web design + Teaching',  
 'BFA Computer Art, Animation and VFX (not fine or performing arts)',  
 'Animal science',  
 'i have not yet attended a university i am still in middle school school',  
 'Home science', 'first chemistry, now computer science',  
 'Costume and fashion designing',  
 'Analise e desenvolvimento de sistemas', 'BA LLB',  
 'Marine Transportation', 'Analytical chemistry',  
 'Estate Management', 'i am a computer science student',  
 'English teacher', 'Computer programing and operations',  
 'Bsc. Physics With Electronics', 'bachelor in arts',  
 'Computer Science, Econometrics', 'Computer application ',  
 'Bio and not going g for a BA in performing arts',  
 'project management specialist',  
 'I have no idea what to choose, I did an intro to C.S Course about a year ago in  
 a university',  
 'Electronics and communication', 'Physics and Electronics ',  
 'currently on a bca course',  
 'I am probably going to do civil engineering', 'Open University',  
 'ICT and Mathematics ', 'Electrical Electronics Engineering',  
 "I don't attend", 'Biomedical engineering', 'Food and beverage',  
 'Marketing', 'computer network engineering',  
 'Currently studying Applied Mathematics and Computer Science',  
 'Finance', 'Education, curriculum design ', 'Library Science',  
 'accounting', 'elementary education', 'Chinese Language Major',  
 'Linguistics ', 'Fire Safety', 'military university',  
 'Currently attending university, (Comp. Sci. Eng.)',  
 'I am in university now',  
 'information software engineering and IT Technician',  
 'English language', 'business',  
 'Accounting and Management Information Systems (comes as a package in my  
 country)',  
 'Communication Disorders', 'Town planning',  
 'Architecture and Urbanism', 'Awaiting admission into University',  
 'Dance Choreography, and Music Design',  
 "I have attended a university in computer game development with a focus on game  
 design, but haven't taken the bachelor's degree. I did however read all courses. ",  
 'Computer Science and English literature and language studies (to become a  
 highschool teacher)',  
 'industrial Technology Education', 'Chemical Engineering',  
 'Entrepreneurship', 'Marine Engineering',

'técnicatura en desarrollo de software', 'Computer Application',  
 'Languages and Translating', "I didn't attend",  
 'Computer Science, Mathematics, Psychology', 'Education degree',  
 'electro-mechanical engineering', 'Culinary',  
 'Currently studying software engineering',  
 'Linguistics and Didactics', 'Data analytics and Management',  
 'City and Regional Planning, Business Management',  
 'Automation and Robotics',  
 'Currently a software development student',  
 'Machine Tool Technology & PLC Programming',  
 'Hotel & Restaurant Management + Clear Professional K-8 Multi-Subject Teaching  
 Credential',  
 'hacking', 'Classical French Baking and Pastry Arts ',  
 'Animal science ', 'Construction',  
 'Film, Business and English (graduate level), Art, Social Work, Geology,  
 Criminal Justice (undergrad)',  
 'Analysis and systems development',  
 'Attended University through a bridging program for 2 years. Bachelor of  
 Arts.',  
 'Aviation maintenance ', 'Systems and informatics engineering',  
 'Mechatronics', 'I do not know whether I will attend a University',  
 'Programación ', 'Public Affairs',  
 'Integrated Science, Business and Technology ',  
 'Architecture Design', "Haven't gotten there yet.",  
 'mantenimiento de equipos de computo',  
 'Criminal justice and intelligence collection ',  
 'Business Administration', 'I am CHILD', 'University Studies',  
 'attending university with computer science major at Uopeople',  
 'business / computer science',  
 'I am currently in high school, soon going to college',  
 'Accounting ', 'Risk Management',  
 'electrical and information technology', 'Spanish',  
 'Have not attended university',  
 'Computer Science Major in Game Dev and Multimedia', 'finance',  
 'Studied Electronics to HNC (first > onc > hnc)', 'Public law',  
 'teaching', 'English literature', 'culinary arts',  
 'Geographical Information Science at University of Edinburgh',  
 'not university, but trade school halfway through an apprenticeship now',  
 'Business Admin', 'Electrical and Electronics',  
 'i chaged three times, i was on mechanical, first then animation, and finally in  
 data engineering',  
 'Information sciences - related to informations (data mining, handling of  
 informations etc.) and librarianship, not so much related with CS-related sciences',  
 'cognitive systems', 'Medical sciences and engineering',  
 'I am still not old enough to attend university',  
 'Currently still in college but thinking of Uni',  
 'european languages', 'Industrial Design Student', 'Urban Studies',  
 'I have not attended a university yet',  
 "I'm still in senior high school",

'have not attended university yet', 'Telecom Engineering',  
 'Agricultural science', 'Contaduría',  
 'project management technology',  
 'Went to college for 6 months before stopping for personal reasons.',  
 'Double Major: Philosophy & Behavioural Science',  
 '\*havent yet attended university ',  
 'I am still studying, in my case Business Administration.',  
 'международная экономика', 'philosophy', 'Graphic Design',  
 'Communications & Spanish', 'not yet attending',  
 "All above. (it would be better if you've used checkboxes)",  
 'Welding ', 'too young', 'Interactive Animation (Game Design ...)',  
 'tecnologia en desarrollo de software',  
 'Ingeniería en Sistemas Computacionales',  
 'Law in the Public Administration', 'Game development',  
 'Arquitectura', 'Currently studying Computer Science & Business',  
 'Human resources', 'Mechatronics engineering',  
 'I started university a moth ago in cyber security',  
 "I'm not in university", 'BARCHELOR OF SCIENCE', 'BCA',  
 'attending college for fine arts/performing arts',  
 "I haven't gone to college yet", 'History', 'Pastoral Ministry',  
 'I am currently close to completing high school',  
 'I am doing diploma in computer engg. after 10th grade currently in 2nd year',  
 'I have not attended university YET',  
 'mathematics, computer science', 'International Relations',  
 'Still in High School', 'Anthropology', 'Microbiology ',  
 'Contador Publico - Faltaron 6 materias para graduarme',  
 'Currently studying computer science', 'liberal arts - general',  
 'Psychology, Accounting', "I'm still in high school.",  
 'Foreign language (Japanese)', 'Culinary arts', 'Filmmaking',  
 'Gastronomy (basically a business course with focus on the restaurant industry)',  
 'Agric Engineering ', 'Safety', 'Electrical Engineering ',  
 'Film Production Technology',  
 'I am still in the university studying computer science.',  
 'psychology', 'Programming foundation ', 'Construction Management',  
 'Medical Microbiology ', 'High School Student', 'Forensic Science',  
 'Técnico em Administração', 'Economy ',  
 'marketing and computer science', 'Gastronomy', 'kinesiology',  
 'Criminalistics',  
 "I'm now attending university, first year computer engineering",  
 'Cyber Security and Media and Communications',  
 'Traduction and Business',  
 'I studied Industrial and Organizational Psychology, so a business discipline with heavy emphasis on the social science aspect of business.',  
 'Seminaries ', 'Attended community college going for AA degree',  
 'Information, IT, Economic',  
 'I was in a hybrid science and engineering program that covers natural and environmental as well as space exploration',  
 'English Language and Literature',

'System Administration, Cyber Security and Information Assurance',  
 'Electronics and communication engineering',  
 'Multimedia & Animation',  
 'I attended to UIT school in Ho Chi Minh City, Vietnam for Computer science. But  
 I have dropped out since the first year.',  
 'Languages ', 'Basic programming in Java',  
 'i did not attend university, Yet.',  
 'Applied mathematics and computer science', 'Criminal', 'Tourism ',  
 "I attended but I didn't finish", 'electronic',  
 'Business discipline and Electronic Engineering',  
 'Artificial Intelligence & Data Science', 'high school',  
 'Currently in University for engineering; this is a hobby of mine.',  
 'Agricultural', 'Architecture and design',  
 'A fake engineer career, lamest education ever, shitty professors, exams too  
 easy and socialism/comunism implemented (UTN argentina) ',  
 'Electrical engineering and Mathematics',  
 'Not yet .. but interested in Computer Science and Software engineering',  
 'Studying diploma in Electrical engineering ',  
 'Exercise physiology', 'Fashion Merchandising', 'Product Design',  
 'Entertainment (3d Film Production)', 'Electrician',  
 'Human Resources', 'Quantity Surveying',  
 'Electronics Engineering', 'Primary School Teacher',  
 'Currently getting my high school diploma.',  
 'haven't attended uni yet', 'Urban Economics',  
 'Digital Media & Design', 'CAD', 'Biology and Computer Science',  
 'Information Studies', 'public administration',  
 'Environmental sciences, Natural sciences, Mathematics',  
 "Haven't attended one yet", 'I am in elementary school',  
 'Aviation Sciences, MBA in Management', 'CSE',  
 'Animal Science and Technology ', 'BTech IT', 'maritime academy',  
 'Public Relations', 'Business statistics economics law',  
 'Economics (unfinished)',  
 'once I attend a university (hopefully full sail) I plan to take computer  
 science',  
 'I am reading in class 8.',  
 'im in secondary school in england year 9', 'Information Design ',  
 'Agribusiness Management', 'Mathematics and then Computer Science',  
 'Mechanical Engineering & MSc Industrial Engineering',  
 "I haven't attended yet.", 'Software Engineering',  
 'Law, Computer Science', 'Master of Computer Application',  
 'Product Design & Quantity Surveying',  
 'I am currently studying electronic engineering.',  
 '3 D graphics ( Autocad ) U.S.G. certificate program 4.0 gpa. ',  
 'I am going to join college',  
 'B.M. in Music Performance, B.S. in Information Technology',  
 'Social Work (this is separate from a social science)',  
 'English and psychology',  
 'I am attending to the university, studying Computer science, computer  
 engineering, or software engineering',

'Secondary student', 'laws ', 'Elementary Education K-8',  
'A business discipline along with computer application',  
'Disaster Preparedness and Environmental Technology',  
'i am istill studying', 'english studies ', 'Tourism',  
'Product and system design',  
'I worked hard in the University of Life.',  
' technical in computer programming ',  
'BA-Journalism, MS- Human-Computer Interaction',  
'accounting, graphic design',  
'I attended several and earned several degrees including a PhD in Physics',  
'Game developing, Informatic Engineer, Applied Languages',  
'Masters in Education', 'pharmacy', 'ElectricalEngineering',  
'Physcial Education', 'Information and Communication Technologies',  
'Modern Applied Languages', 'Criminology and Security studies',  
'I haven't attended university because I can't',  
'Self thought: animation, coding, programming', 'Now starting',  
'translation and interpreting', 'High School Teaching',  
'Artificial Intelligence and Machine Learning (currently in 2nd year)',  
'Geophysics and Mathematics',  
'I am attending now computer science',  
"i studied engineering than biology and now i'm in med school ",  
'im 13 i still cant go to university.', 'Life science',  
'Still on highscool', 'Digital Humanities',  
'Information systems and programming ', 'Biochemistry ',  
'Enginerig with Business', 'commercial photography ',  
'Still in University studying software engineering',  
'Food industry engineer', 'Bioinformatics and Biotechnology',  
'Japanese Language and TESOL', 'Pharmacy ',  
'I am still a student at the university (physiotherapy)',  
'Philology: Language',  
'currently about to start attending law school',  
'undecided and business courses ', "i'm still in university",  
'Mixed between math, physics, computer science, chemistry',  
'going to college this year', 'Veterinary ', 'Geology',  
'Not yet done in high school', 'Economics, Theater',  
"I studied as a manager/entrepreneur of a catering company, but I didn't get my diploma",  
'Marketing management', 'Computer systems Engineering ',  
'not completed - business studies',  
'Currently in first semester of Computer Science degree',  
'Forestry', 'Currently in final year law-LLB',  
'I studied computer science and engineering in a college but it was a waste of time because we were just being taught to pass in exams and no anything about computer science itself.',  
'hospitality management', 'Industrial Engineer',  
'Geography and Environmental Science ',  
'Currently a grade 10 student, so no senior high/college yet',  
'Im in highscool', 'Business', 'Agricultural economics ',  
'Commercial English', 'games and multimedia',

'Health and Safety Technology',  
"I haven't attended university yet",  
'Just starting into computer science',  
'I have a professional qualification in accounting.',  
'Computational Biology', 'Full-Stack Web Dev Bootcamp',  
"I haven't finished highschool yet", 'GE courses',  
'Translation and Interpreting Studies',  
'Economics with Econometrics',  
'Occupational Safety and Health Management ',  
'Still in middle/high school', 'Office Technology and Management',  
'game technology', 'Sound Engineering',  
"I'm about securing admission. Course: Mathematics",  
'In Vietnam, 13 years old is still in grade 7',  
'Planning to Study Mechanical Engineering soon',  
'Undergrad: Public Policy Leadership. Masters: MBA',  
'Integrated Studies (Sustainability and Public Health)',  
'Cybersecurity with mostly business, not sure how to describe',  
'Mathematics', 'allied health science (public health',  
'English language teaching (TEFL)',  
'i havent gone yet but i plan to',  
'I have a degree in Culinary Arts, I was a chef for 25 years, then covid. ',  
'Bible', 'Nursing ', 'Criminal Justice Administration',  
'Industrial pschology',  
'Game Development with Computer Graphics but not finished.',  
'Currently working towards a Digital & Technology Solutions Degree as part of  
the software dev apprenticeship',  
'Hospitality and Tourism Management', 'Consumer Science',  
'Attending a university',  
'BA Ed major in Chinese Language and culture',  
'Technical, Automotive', 'Administración hotelera',  
'Still In School',  
"I haven't finished, but I'm studying data science.",  
'Political science', 'Video Game Design', 'Religious studies',  
'Still in school :)', 'A political sciences discipline ',  
'Computer Science, Mathematics and Data Science ',  
'General Studies: Science', 'law (llb.)', 'Doubled in Criminology',  
'Took a couple nursing classes',  
'Bachelor of Computer Application',  
'Music with a concentration in business', 'Architecture ',  
'reseaux informatique et telecom', 'Information Security',  
'Sports', 'early childhood education',  
'International Global Studies', 'mathematics and CS', 'Nutrition',  
'Currently involved in engineering discipline.(Electronics and Information)',  
'Hospitality Studies', 'TELECOMMUNICATIONS',  
'Mechanical engineering ', 'translator carrear',  
'Administration and business',  
"master's in English and majored in Linguistics", 'religious ',  
'Electrical Engineering',  
'I am only 17 and in year12, I am yet to attend University.'

```
'i am sill studying at a auniversity',
'Currently at computational engineering', 'Bachelor of Arts',
'hotel management', 'Accountancy', 'Engineering',
'First accounting and now IT',
'Mathematics and Secondary Education', 'Hospitality Management',
'English Education', 'college is about to start. cse branch',
'Aun sigo en la universidad', 'nursery teacher',
'finance as bachelor, linguistics as master',
'Agriculture science', 'next year I will go to university',
'Foreign Languages', 'Orthopaedic Surgeon', 'CS-Finance',
'Molecular biology, Biochemistry, Bioinformatics (MB3)',
'Multidisciplinary Studies (nanoscience, psychology, communications)',
'computer information systems', 'Geophysicist',
'Paralegal Studies', 'Multidisciplinary Studies',
'Double major: Computer Information Systems and Finance',
'New Media', 'Applied Computing', 'anthropology',
'Phisical Education'], dtype=object)
```

```
most_common_degrees = (survey_df.UniversityDegree.value_counts(normalize=True).head(15))
most_common_degrees
```

```
I didn't attend a university
22.473182
Computer science, computer engineering, or software engineering
20.530393
Another engineering discipline (e.g., civil, electrical, mechanical)
10.441001
A business discipline (e.g., accounting, finance, marketing)
7.210965
A social science (e.g., sociology, psychology, political science, economics)
5.625745
Information systems, information technology, or system administration
5.148987
A humanities discipline (e.g., literature, history, philosophy)
4.445769
A natural science (e.g., biology, chemistry, physics)
3.933254
Fine arts or performing arts (e.g., graphic design, music, studio, art)
2.949940
A health science (e.g., nursing, pharmacy, radiology)
2.794994
Undecided or no major
1.740167
Mathematics or statistics
1.698451
Communications (e.g., journalism, TV production)
1.293206
Web development or web design
1.203814
Environmental science (e.g., earth sciences, sustainability)
```



0.798570

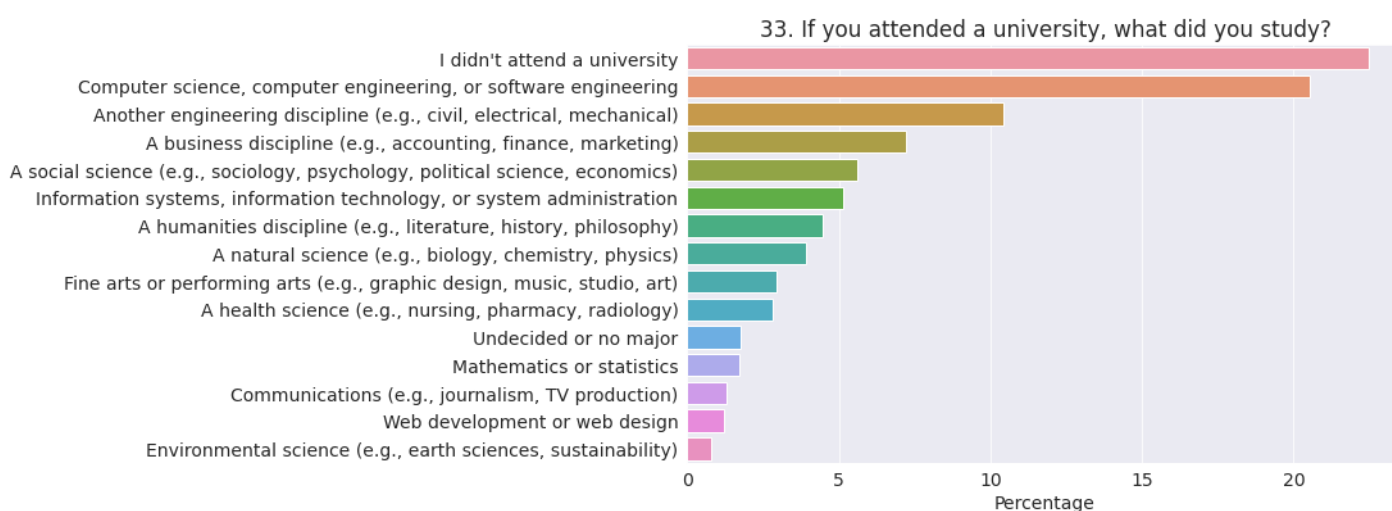
Name: UniversityDegree, dtype: float64

```
plt.figure(figsize=(10,6))
sns.barplot(most_common_degrees, most_common_degrees.index)

plt.title(questions['UniversityDegree'])
plt.ylabel(None)
plt.xlabel('Percentage');
```

/opt/conda/lib/python3.9/site-packages/seaborn/\_decorators.py:36: FutureWarning: Pass the following variables as keyword args: x, y. From version 0.12, the only valid positional argument will be `data`, and passing other arguments without an explicit keyword will result in an error or misinterpretation.

```
warnings.warn(
```



Perhaps surprisingly, the most common answer from respondents was that they did not attend a university, with more than a 1/5th of the respondents saying as much. We will see in the next graph what levels of education are most common among the respondents.

## Highest level of degree achieved by respondents

```
questions['EducationLevel']
```

'32. What is the highest degree or level of school you have completed?'

```
survey_df.EducationLevel.unique()
```

```
array(['Professional degree (MBA, MD, JD, etc.)', 'Bachelor's degree',  
      "Master's degree (non-professional)",  
      'High school diploma or equivalent (GED)',  
      'Some college credit, no degree', nan, 'Some high school',  
      'Associate's degree', 'Trade, technical, or vocational training',  
      'No high school (secondary school)', 'Ph.D.'], dtype=object)
```

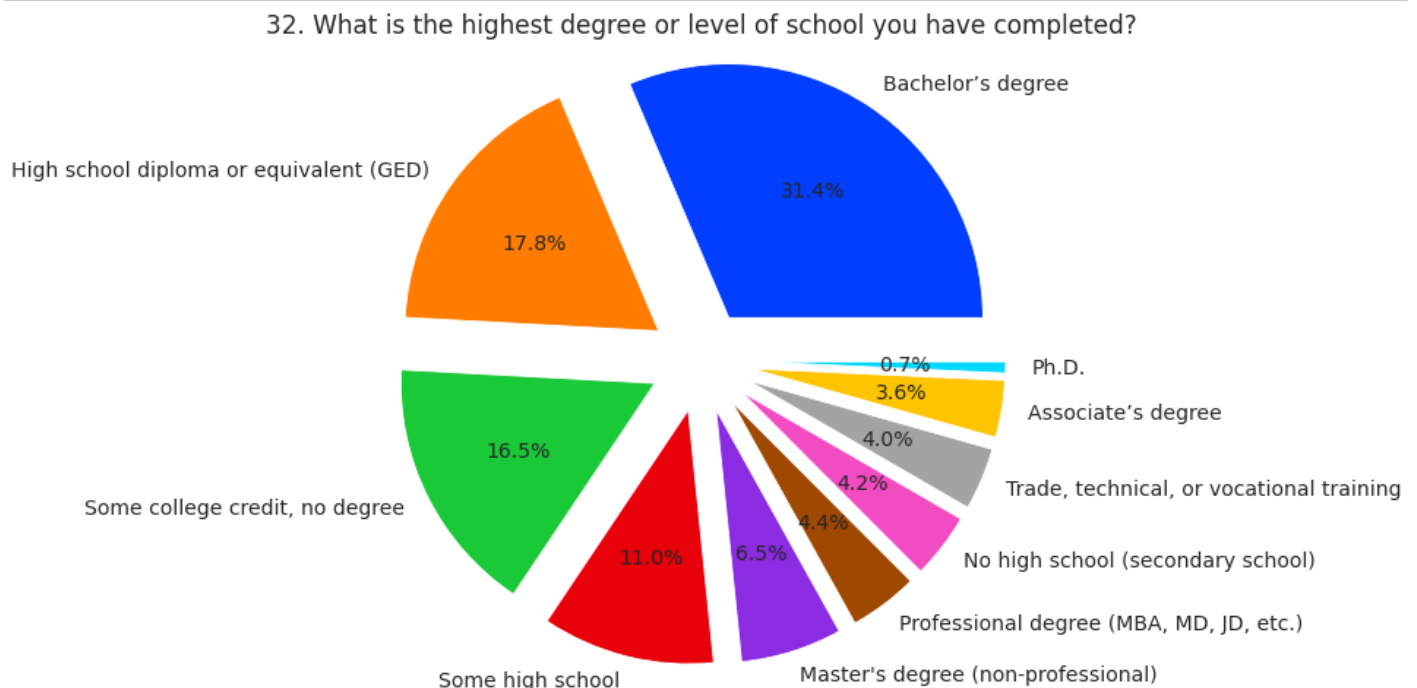
```
highest_degree_level = survey_df.EducationLevel.value_counts()  
highest_degree_level
```

Bachelor's degree	5340
High school diploma or equivalent (GED)	3024
Some college credit, no degree	2800
Some high school	1872
Master's degree (non-professional)	1101
Professional degree (MBA, MD, JD, etc.)	750
No high school (secondary school)	709
Trade, technical, or vocational training	677
Associate's degree	619
Ph.D.	125

Name: EducationLevel, dtype: int64

We will use a pie chart to visualise the results. As we have a lot of options here, we will separate out our chart using the 'explode' feature

```
plt.figure(figsize=(8,8))
colors = sns.color_palette('bright')
plt.pie(highest_degree_level,
        labels=highest_degree_level.index,
        colors=colors,
        autopct='%.1f%%',
        explode = (0.2,0.2,0.2,0.2,0.2,0.2,0.2,0.2,0.2,0.2))
plt.title(questions['EducationLevel']);
```



The most common level of degree attained by the respondents was a bachelor's degree, most likely in computing science as shown in the previous bar graph highlighting what courses were studied the most at university. The previous 2 graphs highlight that there is no one set path to becoming a coder, every individual has their own path which can work as long as they are willing to put in the effort to learn.

Let us save and upload our work to Jovian before continuing

```
import jovian
```

```
jovian.commit()
```

```
[jovian] Updating notebook "adammasongh/2021-new-coder-survey" on https://jovian.com  
[jovian] Committed successfully! https://jovian.com/adammasongh/2021-new-coder-survey  
'https://jovian.com/adammasongh/2021-new-coder-survey'
```

## Asking and Answering Questions

Now that we have learned more about the respondent's background, we can begin asking specific questions to gain more insight into the dataset. We will attempt to answer 5 questions which will hopefully help us understand more about the respondent's thoughts and opinions on coding.

### Q: What online resources are used most by new coders?

To start, we will look at the list of columns and associated questions to see which columns we can use to answer this.

```
questions
```

Column

BiggestReasonToCode	1. What is your biggest reason for learning to...
MethodsUsedToLearn	2. What methods have you used to learn about c...
OnlineResourcesUsed	3. Which online learning resources have you fo...
UsefulCodingEvents	4. If you have attended in-person coding-relat...
UsefulCodingPodcasts	5. If you have listened to coding-related podc...
UsefulCodingVideos	6. If you have watched coding-related YouTube ...
HoursSpentLearningPerWeek	7. About how many hours do you spend learning ...
MonthsSpentProgramming	8. About how many months have you been program...
MoneySpentLearning	9. Aside from university tuition, about how mu...
AlreadyEmployed	10. Are you already employed in a software dev...
FirstSoftwareJob	11. If you are already employed as a developer...
InterestedInSoftware	12. If you are NOT already a developer, are yo...
PreferredCareerPath	13. If you are interested in a software develo...
CareersInterestedIn	14. Which of these careers are you interested in?
WhenApplyingForJobs	15. When do you plan to start applying for dev...
ExpectedSalary	16. About how much money do you expect to earn...
ReasonsForSoftwareInterest	17. Please select up to 3 reasons why you are ...
WorkFromHomeVsOffice	18. After the pandemic, how many days would yo...
WillingToRelocate	19. Are you willing to relocate for a job?
EmploymentStatus	20. Regarding employment status, are you curre...
CurrentFieldOfWork	21. If you are currently working, which field ...
MoneyMadeLastYear	22. About how much money did you earn last yea...
Age	23. How old are you?
Gender	24. Which of the following best represents how...
Ethnicity	25. With which of these groups do you primaril...
Region	26. Which part of the world do you live in?
State	27. If you are living in the US, which state d...
EducationLevel	32. What is the highest degree or level of sch...
UniversityDegree	33. If you attended a university, what did you...

CurrentStudent	34. Are you currently a student attending or e...
TimeSpentInCurrentRole	40. How long have you been working in your cur...
MonthsSpentLookingForJob	41. Before you got your last job, how many mon...
LikelyToLoseJob	42. If you are working, thinking about the nex...
FindNewJobEasiness	43. If you are working, how easy would it be f...
ConsideredUnderEmployed	44. Do you consider yourself under-employed? (...)
EarningSatisfaction	45. Please tell us how satisfied you are with ...
BenefitsSatisfaction	45. Please tell us how satisfied you are with ...
JobSecuritySatisfaction	45. Please tell us how satisfied you are with ...
WorkLifeBalanceSatisfaction	45. Please tell us how satisfied you are with ...
GrowthOpportunitySatisfaction	45. Please tell us how satisfied you are with ...
WorkCultureSatisfaction	45. Please tell us how satisfied you are with ...
WorkDiversitySatisfaction	45. Please tell us how satisfied you are with ...
WorkloadSatisfaction	45. Please tell us how satisfied you are with ...

Name: QuestionText, dtype: object

The 'MethodsUsedToLearn', 'OnlineResourcesUsed', 'UsefulCodingEvents', 'UsefulCodingPodcasts' and 'UsefulCodingVideos' seem to be the most suitable columns to answer this.

We will create a sub-dataset containing the info from the columns specified above.

```
resource_columns = ['MethodsUsedToLearn',
                    'OnlineResourcesUsed',
                    'UsefulCodingEvents',
                    'UsefulCodingPodcasts',
                    'UsefulCodingVideos']
```

```
resources_df = survey_df[resource_columns]
resources_df
```

	MethodsUsedToLearn	OnlineResourcesUsed	UsefulCodingEvents	UsefulCodingPodcasts	UsefulCodingVideos
0	Online resources, Books, In-person bootcamps, ...	freeCodeCamp, Mozilla Developer Network (MDN),...	conferences, workshops, Meetup.com events	The Changelog	CS Dojo, freeCodeCamp
1	Online resources, Books, Online bootcamps	freeCodeCamp, Mozilla Developer Network (MDN),...	I haven't attended any in-person coding-relate...	The Changelog, Code Newbie Podcast	Adrian Twarog, Code with Ania Kubów, Coder Cod...
2	Online resources, Books, Hackathons, Meetup.co...	freeCodeCamp, Mozilla Developer Network (MDN),...	Meetup.com events	I haven't listened to any podcasts	AmigosCode, Dev Ed, freeCodeCamp, Kevin Powell...
3	Online resources, Books	freeCodeCamp, Mozilla Developer Network (MDN),...	I haven't attended any in-person coding-relate...	Darknet Diaries, Real Python Podcast	freeCodeCamp, Traversy Media
4	Online resources, Books, Online bootcamps	freeCodeCamp, Stack Overflow, Coursera, Udemy	I haven't attended any in-person coding-relate...	Talk Python to Me	freeCodeCamp, The Net Ninja, Traversy Media
...	...	...	...	...	...
18121	Online resources, Books, In-person bootcamps	Stack Overflow, EdX, Khan Academy, Udacity, Ci...	conferences, workshops	Le code a changé (RadioFrance app)	Hashtag Programação, Curso em Video, Boson Tre...

	MethodsUsedToLearn	OnlineResourcesUsed	UsefulCodingEvents	UsefulCodingPodcasts	UsefulCodingVideos
18122	Online resources, In-person bootcamps	freeCodeCamp, Mozilla Developer Network (MDN),...	I haven't attended any in-person coding-relate...	I haven't listened to any podcasts	freeCodeCamp
18123	Online resources, Books	freeCodeCamp, Mozilla Developer Network (MDN),...	I haven't attended any in-person coding-relate...	I haven't listened to any podcasts	Ben Awad, Coder Coder, Coding Train, Dev Ed, f...
18124	Online resources, Books, In-person bootcamps, ...	freeCodeCamp, Stack Overflow, Codecademy, Udem...	I haven't attended any in-person coding-relate...	I haven't listened to any podcasts	AmigosCode, Code with Ania Kubów, freeCodeCamp...
18125	Online resources, Books	freeCodeCamp, Mozilla Developer Network (MDN),...	I haven't attended any in-person coding-relate...	I haven't listened to any podcasts	Code with Ania Kubów, Coding Train, Dev Ed, fr...

17938 rows × 5 columns

```
resources_df.MethodsUsedToLearn.nunique()
```

1718

Some of these columns allow for multiple answers from the user, which is why columns such as MethodsUsedToLearn has over 1700 unique answers. It would be useful to split up each answer that has been given so we can get a better idea of the most used resources. We will use the `split_multicolumn()` function which has been taken from Lesson 6 of the Data Analysis with Python: Zero to Pandas course by Jovian and slightly modified (<https://jovian.com/learn/data-analysis-with-python-zero-to-pandas/lesson/lesson-6-exploratory-data-analysis-a-case-study>).

```
def split_multicolumn(col_series):
    result_df = col_series.to_frame()
    options = []
    # Iterate over the column
    for idx, value in col_series[col_series.notnull()].iteritems():
        # Break each value into list of options
        for option in value.split(','):
            option=option.strip() # removes whitespace to ensure no duplicate columns
            # Add the option as a column to result
            if not option in result_df.columns:
                options.append(option)
                result_df[option] = False
            # Mark the value in the option column as True
            result_df.at[idx, option] = True
    return result_df[options]
```

```
resource_type_df = split_multicolumn(resources_df.MethodsUsedToLearn)
resource_type_df
```

/tmp/ipykernel\_56/226654725.py:12: PerformanceWarning: DataFrame is highly fragmented. This is usually the result of calling `frame.insert` many times, which has poor performance. Consider joining all columns at once using `pd.concat(axis=1)` instead. To get a de-fragmented frame, use ``newframe = frame.copy()``

```
result_df[option] = False
```

	Online resources	Books	In-person bootcamps	Online bootcamps	Meetup.com events	Conferences	Workshops	Hackathons	Tinkering	Uni
0	True	True	True	True	True	True	True	False	False	
1	True	True	False	True	False	False	False	False	False	
2	True	True	False	False	True	False	True	True	False	
3	True	True	False	False	False	False	False	False	False	
4	True	True	False	True	False	False	False	False	False	
...	...	...	...	...	...	...	...	...	...	...
18121	True	True	True	False	False	False	False	False	False	
18122	True	False	True	False	False	False	False	False	False	
18123	True	True	False	False	False	False	False	False	False	
18124	True	True	True	True	False	False	False	False	False	
18125	True	True	False	False	False	False	False	False	False	

17938 rows × 1405 columns

With this information, we can have a look at the 10 most common methods used by new coders to learn coding.

```
resource_type_totals = resource_type_df.sum().sort_values(ascending=False)
resource_type_totals
```

```
Online resources          16857
Books                    7682
Online bootcamps         4583
Workshops                1757
In-person bootcamps       950
...
I'm Completely New!      1
sharepoint designer and now powerapps (at work) 1
infopath                 1
youtube videos and school 1
in-person trainings       1
Length: 1405, dtype: int64
```

```
resource_type_pct = resource_type_totals.head(10) * 100 / len(resource_type_df)
resource_type_pct
```

```
Online resources    93.973687
Books               42.825287
Online bootcamps    25.549114
Workshops           9.794849
In-person bootcamps 5.296020
Conferences          4.777567
Hackathons           4.588025
Meetup.com events   3.066117
```

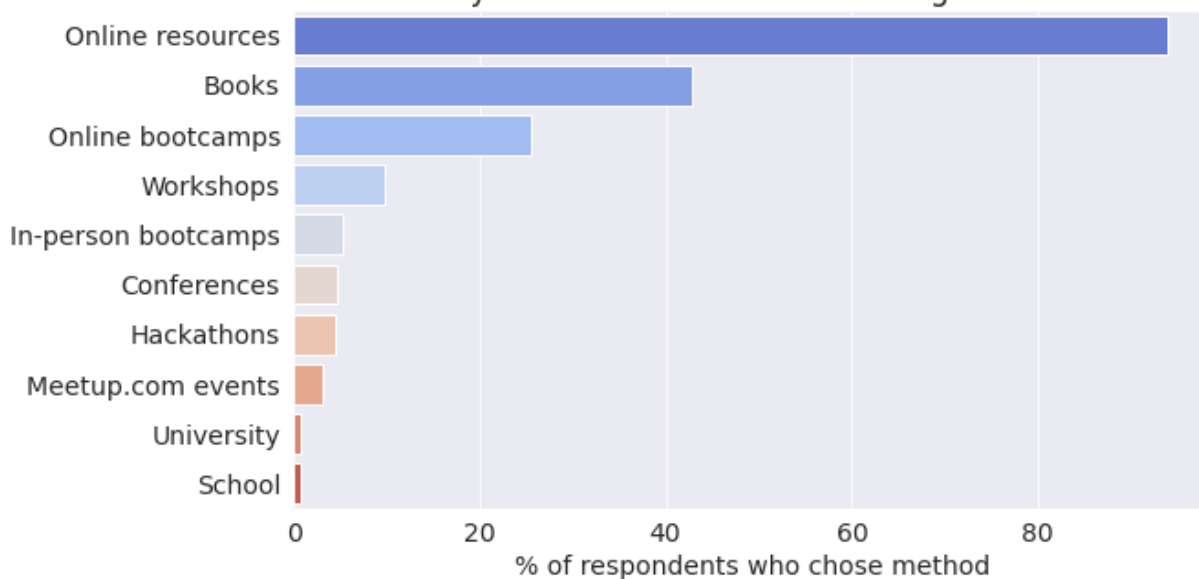
```
University          0.741443
School              0.724718
dtype: float64
```

```
sns.barplot(resource_type_pct, resource_type_pct.index, palette='coolwarm')
plt.title(questions['MethodsUsedToLearn'])
plt.ylabel(None)
plt.xlabel('% of respondents who chose method');
```

/opt/conda/lib/python3.9/site-packages/seaborn/\_decorators.py:36: FutureWarning: Pass the following variables as keyword args: x, y. From version 0.12, the only valid positional argument will be `data`, and passing other arguments without an explicit keyword will result in an error or misinterpretation.

```
warnings.warn(
```

## 2. What methods have you used to learn about coding? Please select all that apply.



With over 90% of respondents saying yes, it should come as no surprise that online resources are the most commonly used methods that coders use to learn their programming skills. This can be attributed to the volume of content readily available (usually free of cost!), as well as the ability to pick and choose the specific topic you wish to learn about. Books were also a fairly common source of learning among respondents, with over 40% claiming they read books to learn more about coding.

## Q: Do men and women have different habits and opinions regarding coding?

We know from our previous analysis that the majority of respondents were male. It would be interesting to know whether there are differences to how men and women approach and think about coding. Whilst there were respondent's who chose other genders, we are only focusing on men and women here as the sample size for other options was too low.

We will start by creating two separate dataframes containing only male and female respondents. We will then join them together using the `.concat()` function.

```
male_df = survey_df[survey_df["Gender"] == 'Male']
male_df
```

	BiggestReasonToCode	MethodsUsedToLearn	OnlineResourcesUsed	UsefulCodingEvents	UsefulCodingPodcasts
1	To change careers	Online resources, Books, Online bootcamps	freeCodeCamp, Mozilla Developer Network (MDN),...	I haven't attended any in-person coding-relate...	The Changelog, Code Newbie Podcast
2	To change careers	Online resources, Books, Hackathons, Meetup.co...	freeCodeCamp, Mozilla Developer Network (MDN),...	Meetup.com events	I haven't listened to any podcasts
3	As a hobby	Online resources, Books	freeCodeCamp, Mozilla Developer Network (MDN),...	I haven't attended any in-person coding-relate...	Darknet Diaries, Real Python Podcast
4	To start your first career	Online resources, Books, Online bootcamps	freeCodeCamp, Stack Overflow, Coursera, Udemy	I haven't attended any in-person coding-relate...	Talk Python to Me
5	To succeed in current career	Online resources, Books, Conferences	freeCodeCamp, Mozilla Developer Network (MDN),...	I haven't attended any in-person coding-relate...	Ladybug Podcast
...	...	...	...	...	...
18120	To change careers	Online resources, In-person bootcamps, Online ...	freeCodeCamp, Mozilla Developer Network (MDN),...	I haven't attended any in-person coding-relate...	Hipsters Ponto Tech ( a brazilian podcast )
18121	To change careers	Online resources, Books, In-person bootcamps	Stack Overflow, EdX, Khan Academy, Udacity, Ci...	conferences, workshops	Le code a changé (RadioFrance app)
18123	To start your first career	Online resources, Books	freeCodeCamp, Mozilla Developer Network (MDN),...	I haven't attended any in-person coding-relate...	I haven't listened to any podcasts
18124	To change careers	Online resources, Books, In-person bootcamps, ...	freeCodeCamp, Stack Overflow, Codecademy, Udem...	I haven't attended any in-person coding-relate...	I haven't listened to any podcasts
18125	To succeed in current career	Online resources, Books	freeCodeCamp, Mozilla Developer Network (MDN),...	I haven't attended any in-person coding-relate...	I haven't listened to any podcasts

12074 rows × 43 columns

```
female_df = survey_df[survey_df['Gender'] == 'Female']
female_df
```

	BiggestReasonToCode	MethodsUsedToLearn	OnlineResourcesUsed	UsefulCodingEvents	UsefulCodingPodcasts
24	To succeed in current career	Online resources, Hackathons, Meetup.com events	freeCodeCamp, Mozilla Developer Network (MDN),...	I haven't attended any in-person coding-relate...	The Changelog, Code Newbie Podcast, Ladybug Po...
40	To change careers	Online resources, In-person bootcamps	freeCodeCamp, Mozilla Developer Network (MDN),...	I haven't attended any in-person coding-relate...	I haven't listened to any podcasts
50	To change careers	Online resources, In-person bootcamps, Online ...	freeCodeCamp, Mozilla Developer Network (MDN),...	I haven't attended any in-person coding-relate...	NaN
69	To start your first career	Online resources	freeCodeCamp, Mozilla Developer Network (MDN),...	I haven't attended any in-person coding-relate...	Syntax.fm, The CSS Podcast
79	To change careers	Online resources, Books, Online bootcamps, Wor...	freeCodeCamp, Mozilla Developer Network (MDN),...	I haven't attended any in-person coding-relate...	JS PARTY



	BiggestReasonToCode	MethodsUsedToLearn	OnlineResourcesUsed	UsefulCodingEvents	UsefulCodingPodcasts
...	...	...	...	...	...
18106	To change careers	Online resources, In-person bootcamps, Online ...	Mozilla Developer Network (MDN), Stack Overflo...	I haven't attended any in-person coding-relate...	I haven't listened to any podcasts
18110	To meet school requirements	Online resources	EdX, Coursera, Khan Academy, Codecademy	I haven't attended any in-person coding-relate...	I haven't listened to any podcasts
18113	To start your first career	Online resources, Books	freeCodeCamp, Stack Overflow, Khan Academy, CS...	I haven't attended any in-person coding-relate...	I haven't listened to any podcasts
18116	To start your first career	Online resources	freeCodeCamp, Codecademy	I haven't attended any in-person coding-relate...	I haven't listened to any podcasts
18122	To change careers	Online resources, In-person bootcamps	freeCodeCamp, Mozilla Developer Network (MDN),...	I haven't attended any in-person coding-relate...	I haven't listened to any podcasts

4456 rows × 43 columns

```
male_female_df = pd.concat([male_df, female_df])
male_female_df
```

	BiggestReasonToCode	MethodsUsedToLearn	OnlineResourcesUsed	UsefulCodingEvents	UsefulCodingPodcasts
1	To change careers	Online resources, Books, Online bootcamps	freeCodeCamp, Mozilla Developer Network (MDN),...	I haven't attended any in-person coding-relate...	The Changelog, Code Newbie Podcast
2	To change careers	Online resources, Books, Hackathons, Meetup.co...	freeCodeCamp, Mozilla Developer Network (MDN),...	Meetup.com events	I haven't listened to any podcasts
3	As a hobby	Online resources, Books	freeCodeCamp, Mozilla Developer Network (MDN),...	I haven't attended any in-person coding-relate...	Darknet Diaries, Real Python Podcast
4	To start your first career	Online resources, Books, Online bootcamps	freeCodeCamp, Stack Overflow, Coursera, Udemy	I haven't attended any in-person coding-relate...	Talk Python to Me
5	To succeed in current career	Online resources, Books, Conferences	freeCodeCamp, Mozilla Developer Network (MDN),...	I haven't attended any in-person coding-relate...	Ladybug Podcast
...	...	...	...	...	...
18106	To change careers	Online resources, In-person bootcamps, Online ...	Mozilla Developer Network (MDN), Stack Overflo...	I haven't attended any in-person coding-relate...	I haven't listened to any podcasts
18110	To meet school requirements	Online resources	EdX, Coursera, Khan Academy, Codecademy	I haven't attended any in-person coding-relate...	I haven't listened to any podcasts
18113	To start your first career	Online resources, Books	freeCodeCamp, Stack Overflow, Khan Academy, CS...	I haven't attended any in-person coding-relate...	I haven't listened to any podcasts
18116	To start your first career	Online resources	freeCodeCamp, Codecademy	I haven't attended any in-person coding-relate...	I haven't listened to any podcasts
18122	To change careers	Online resources, In-person bootcamps	freeCodeCamp, Mozilla Developer Network (MDN),...	I haven't attended any in-person coding-relate...	I haven't listened to any podcasts

16530 rows × 43 columns

```
male_female_df.Gender.unique()
```

```
array(['Male', 'Female'], dtype=object)
```

We will focus on 6 aspects about coding to see if there are any differences between the two genders. We will use the .groupby() method to get the average for each column of values for both males and females.

```
gender_diff_df = male_female_df.groupby('Gender')[['HoursSpentLearningPerWeek', 'MonthsSpentProgramming', 'MoneySpentLearning', 'Age', 'LikelyToLoseJob', 'F']]
gender_diff_df
```

	HoursSpentLearningPerWeek	MonthsSpentProgramming	MoneySpentLearning	Age	LikelyToLoseJob	F
Gender						
Female	10.735281	10.581227	593.869810	27.616459	2.926064	
Male	13.372549	17.765298	584.484911	26.874541	2.984310	

We will use a subplot to visualise the results.

```
fig, axes = plt.subplots(2,3, figsize = (24,16))

axes[0,0].set_title('Hours Spent Learning Each Week')
sns.barplot(gender_diff_df.index, gender_diff_df.HoursSpentLearningPerWeek,
            ax=axes[0,0])
axes[0,0].set_ylabel('Hours spent');

axes[0,1].set_title('Months Spent Programming')
sns.barplot(gender_diff_df.index, gender_diff_df.MonthsSpentProgramming,
            ax=axes[0,1])
axes[0,1].set_ylabel('Months spent');

axes[0,2].set_title('Money Spent Learning')
sns.barplot(gender_diff_df.index, gender_diff_df.MoneySpentLearning,
            ax=axes[0,2])
axes[0,2].set_ylabel('Money spent');

genders = male_female_df.Gender.unique()
axes[1,0].hist([male_female_df.loc[male_female_df.Gender == x, 'Age'] for x in genders],
               bins=np.arange(10,90,5))
axes[1,0].set_title('How old are you?')
axes[1,0].legend(['Male', 'Female'])
axes[1,0].set_ylabel('Number of respondents')
axes[1,0].set_xlabel('Age');

axes[1,1].set_title('Likely To Lose Job?')
sns.barplot(gender_diff_df.index, gender_diff_df.LikelyToLoseJob,
```

```

        ax=axes[1,1])
axes[1,1].set_ylabel('Likelihood')
axes[1,1].set_yticks(np.arange(0,11,1));

axes[1,2].set_title('Likely To Find new job?')
sns.barplot(gender_diff_df.index, gender_diff_df.FindNewJobEasiness,
            ax=axes[1,2])
axes[1,2].set_ylabel('Likelihood')
axes[1,2].set_yticks(np.arange(0,11,1));

plt.tight_layout(pad=2);

```

/opt/conda/lib/python3.9/site-packages/seaborn/\_decorators.py:36: FutureWarning: Pass the following variables as keyword args: x, y. From version 0.12, the only valid positional argument will be `data`, and passing other arguments without an explicit keyword will result in an error or misinterpretation.

```
warnings.warn(
```

/opt/conda/lib/python3.9/site-packages/seaborn/\_decorators.py:36: FutureWarning: Pass the following variables as keyword args: x, y. From version 0.12, the only valid positional argument will be `data`, and passing other arguments without an explicit keyword will result in an error or misinterpretation.

```
warnings.warn(
```

/opt/conda/lib/python3.9/site-packages/seaborn/\_decorators.py:36: FutureWarning: Pass the following variables as keyword args: x, y. From version 0.12, the only valid positional argument will be `data`, and passing other arguments without an explicit keyword will result in an error or misinterpretation.

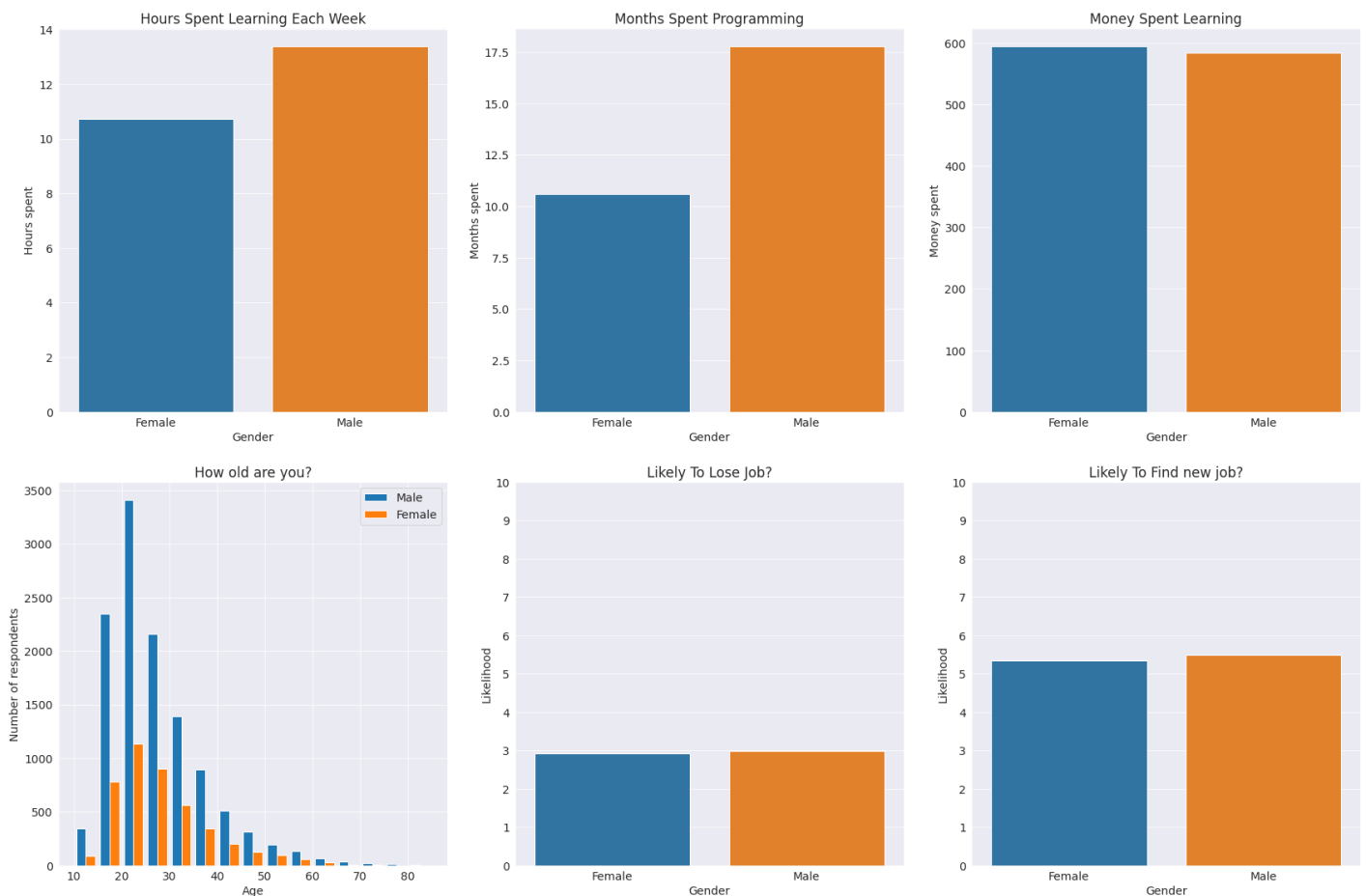
```
warnings.warn(
```

/opt/conda/lib/python3.9/site-packages/seaborn/\_decorators.py:36: FutureWarning: Pass the following variables as keyword args: x, y. From version 0.12, the only valid positional argument will be `data`, and passing other arguments without an explicit keyword will result in an error or misinterpretation.

```
warnings.warn(
```

/opt/conda/lib/python3.9/site-packages/seaborn/\_decorators.py:36: FutureWarning: Pass the following variables as keyword args: x, y. From version 0.12, the only valid positional argument will be `data`, and passing other arguments without an explicit keyword will result in an error or misinterpretation.

```
warnings.warn(
```



From the results, it appears that men spend longer learning how to code each week and have spent more months programming than women on average. There's no significant difference when it comes to money spent learning and the age distribution curve for both appears to follow a similar trend, with 20-40 being where the majority of respondents fall under. Both men and women are fairly confident they won't lose their job anytime soon, and both are unsure whether they are likely to find a new job within a few months if they were fired.

## Q: Do different ethnic groups have different habits and opinions regarding coding?

We now want to see if there are any differences depending on the ethnicity of the respondents. We will use the same approach as we used in the last section.

```
survey_df.Ethnicity.value_counts().head(10)
```

White	6162
South Asian	2434
Hispanic/Latino(a/x)	2425
Black/African American	1938
Southeast Asian	1605
Middle Eastern or North African	678
East Asian	514
Indian	133
African	118
American Indian or Alaska Native	51

Name: Ethnicity, dtype: int64

```

white_df = survey_df[survey_df["Ethnicity"] == 'White']
black_df = survey_df[survey_df["Ethnicity"] == 'Black/African American']
southasian_df = survey_df[survey_df["Ethnicity"] == 'South Asian']
southeastasian_df = survey_df[survey_df["Ethnicity"] == 'Southeast Asian']
eastasian_df = survey_df[survey_df["Ethnicity"] == 'East Asian']
middleeast_df = survey_df[survey_df["Ethnicity"] == 'Middle Eastern or North African']
hispanic_df = survey_df[survey_df["Ethnicity"] == 'Hispanic/Latino(a/x)']

```

```

ethnicity_df = pd.concat([white_df, black_df, southasian_df, southeastasian_df, eastasi
ethnicity_df

```

	BiggestReasonToCode	MethodsUsedToLearn	OnlineResourcesUsed	UsefulCodingEvents	UsefulCodingPodcasts
7	As a hobby	Online resources, Hackathons, Workshops, Tinke...	freeCodeCamp, Mozilla Developer Network (MDN),...	I haven't attended any in-person coding-relate...	The Changelog, Indie Hackers
8	To change careers	Online resources, Books, Online bootcamps	freeCodeCamp, Mozilla Developer Network (MDN),...	I haven't attended any in-person coding-relate...	The CSS Podcast
10	To change careers	Online resources, Books	freeCodeCamp, Stack Overflow, EdX, Khan Academ...	conferences	Code Newbie Podcast, Practical AI, Software En...
14	To succeed in current career	Online resources, Online bootcamps	freeCodeCamp, Mozilla Developer Network (MDN),...	I haven't attended any in-person coding-relate...	I haven't listened to any podcasts
22	Good question	Online resources, Books	freeCodeCamp, Mozilla Developer Network (MDN),...	I haven't attended any in-person coding-relate...	PyBites Podcast
...	...	...	...	...	...
18071	To start your first career	Online resources, Books, School	Mozilla Developer Network (MDN), Stack Overflo...	I haven't attended any in-person coding-relate...	Code Newbie Podcast
18080	To succeed in current career	Online resources, High school and College	freeCodeCamp, Coursera, Codecademy	I haven't attended any in-person coding-relate...	I haven't listened to any podcasts
18085	To change careers	Online resources	Udemy, youtube	I haven't attended any in-person coding-relate...	I haven't listened to any podcasts
18092	To create art or entertainment	Online resources	freeCodeCamp	I haven't attended any in-person coding-relate...	I haven't listened to any podcasts
18109	To succeed in current career	Online resources, Books	freeCodeCamp, Stack Overflow, Udemy	I haven't attended any in-person coding-relate...	I haven't listened to any podcasts

15756 rows × 43 columns

```

ethnicity_diff_df = ethnicity_df.groupby("Ethnicity")[['HoursSpentLearningPerWeek', 'Mor
ethnicity_diff_df

```

	HoursSpentLearningPerWeek	MonthsSpentProgramming	MoneySpentLearning	Age	Likely1
Ethnicity					

	HoursSpentLearningPerWeek	MonthsSpentProgramming	MoneySpentLearning	Age	Likely1
Ethnicity					
Black/African American	13.381686	8.215309	362.835397	26.775936	
East Asian	12.045738	12.593814	775.105641	25.820926	
Hispanic/Latino(a/x)	11.911923	16.686793	387.430877	27.745508	
Middle Eastern or North African	15.651592	12.933408	158.401872	23.577844	
South Asian	14.567095	14.995504	493.195098	22.208522	
Southeast Asian	13.122706	10.891089	1974.176589	22.933847	
White	11.411872	20.432481	496.763420	30.361014	

```
fig, axes = plt.subplots(2,3, figsize = (20,12))

axes[0,0].set_title('Hours Spent Learning Each Week')
sns.barplot(ethnicity_diff_df.HoursSpentLearningPerWeek, ethnicity_diff_df.index,
            ax=axes[0,0])
axes[0,0].set_ylabel(None)
axes[0,0].set_xlabel('Hours spent');

axes[0,1].set_title('Months Spent Programming')
sns.barplot(ethnicity_diff_df.MonthsSpentProgramming, ethnicity_diff_df.index,
            ax=axes[0,1])
axes[0,1].set_ylabel(None)
axes[0,1].set_xlabel('Months spent');

axes[0,2].set_title('Money Spent Learning')
sns.barplot(ethnicity_diff_df.MoneySpentLearning, ethnicity_diff_df.index,
            ax=axes[0,2])
axes[0,2].set_ylabel(None)
axes[0,2].set_xlabel('Money spent');

axes[1,0].set_title('How old are you?')
sns.barplot(ethnicity_diff_df.Age, ethnicity_diff_df.index,
            ax=axes[1,0])
axes[1,0].set_ylabel(None)
axes[1,0].set_xlabel('Age');

axes[1,1].set_title('Likely To Lose Job?')
sns.barplot(ethnicity_diff_df.LikelyToLoseJob, ethnicity_diff_df.index,
            ax=axes[1,1])
axes[1,1].set_ylabel(None)
axes[1,1].set_xlabel('Likelihood')
axes[1,1].set_xticks(np.arange(0,11,1));
```

```
axes[1,2].set_title('Likely To Find new job?')
sns.barplot(ethnicity_diff_df.FindNewJobEasiness, ethnicity_diff_df.index,
            ax=axes[1,2])
axes[1,2].set_ylabel(None)
axes[1,2].set_xlabel('Likelihood')
axes[1,2].set_xticks(np.arange(0,11,1));

plt.tight_layout(pad=2);
```

/opt/conda/lib/python3.9/site-packages/seaborn/\_decorators.py:36: FutureWarning: Pass the following variables as keyword args: x, y. From version 0.12, the only valid positional argument will be `data`, and passing other arguments without an explicit keyword will result in an error or misinterpretation.

```
warnings.warn(
```

/opt/conda/lib/python3.9/site-packages/seaborn/\_decorators.py:36: FutureWarning: Pass the following variables as keyword args: x, y. From version 0.12, the only valid positional argument will be `data`, and passing other arguments without an explicit keyword will result in an error or misinterpretation.

```
warnings.warn(
```

/opt/conda/lib/python3.9/site-packages/seaborn/\_decorators.py:36: FutureWarning: Pass the following variables as keyword args: x, y. From version 0.12, the only valid positional argument will be `data`, and passing other arguments without an explicit keyword will result in an error or misinterpretation.

```
warnings.warn(
```

/opt/conda/lib/python3.9/site-packages/seaborn/\_decorators.py:36: FutureWarning: Pass the following variables as keyword args: x, y. From version 0.12, the only valid positional argument will be `data`, and passing other arguments without an explicit keyword will result in an error or misinterpretation.

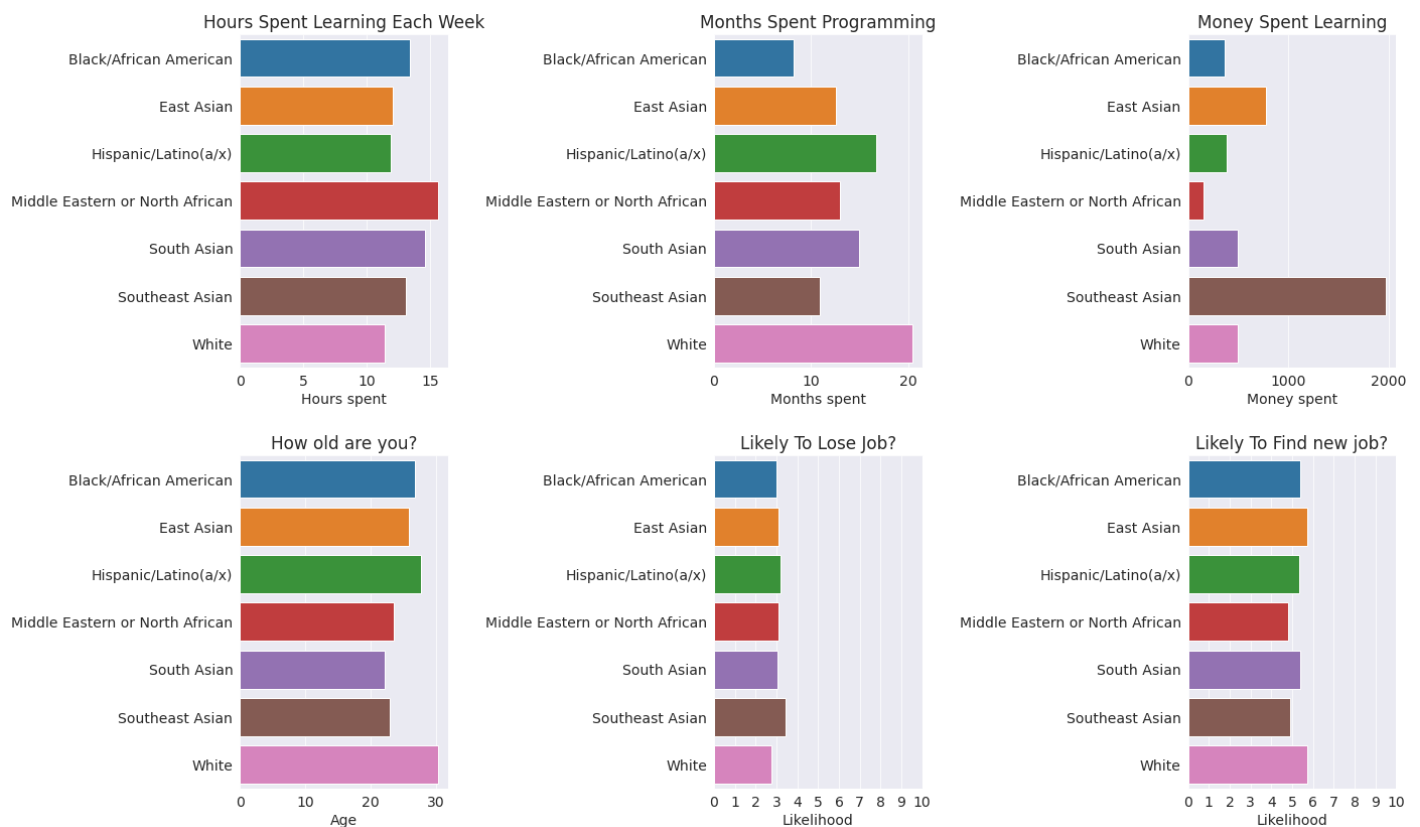
```
warnings.warn(
```

/opt/conda/lib/python3.9/site-packages/seaborn/\_decorators.py:36: FutureWarning: Pass the following variables as keyword args: x, y. From version 0.12, the only valid positional argument will be `data`, and passing other arguments without an explicit keyword will result in an error or misinterpretation.

```
warnings.warn(
```

/opt/conda/lib/python3.9/site-packages/seaborn/\_decorators.py:36: FutureWarning: Pass the following variables as keyword args: x, y. From version 0.12, the only valid positional argument will be `data`, and passing other arguments without an explicit keyword will result in an error or misinterpretation.

```
warnings.warn(
```



In terms of hours spent each week learning, Middle Eastern or North African respondents appear to spend the most hours learning, followed closely by South Asians. East Asian, Hispanic/Latino and White respondents spent the least time learning each week.

White respondents have spent more months programming than any other ethnicity group, with Black/African Americans having spent the least. This could be explained by the fact that white respondents were on average older than any other ethnic group, indicating they have had more time to learn how to program.

Southeast Asians respondents spent by far the most money learning how to code with the average respondent claiming to have spent almost \$2000, nearly 3 times more than South Asians who spend the second most amount of money. Middle Eastern and North African respondents spent the least amount of money learning to code.

All ethnic groups feel it is unlikely they will lose their jobs within the next year, which is encouraging to see, and all groups are unsure whether they are likely to find a new job within the next year if they lost their current job, with Middle Eastern and North African respondents being slightly less confident than the other groups.

## Q: How satisfied are developers in their current roles?

Question 45 in the dataset asks about how satisfied coders are with various aspects of their job. We can use the results to paint a better idea of what some of the main pros and cons are of working within a development environment.

```
satisfaction_columns = ['EarningSatisfaction',
                        'BenefitsSatisfaction',
                        'JobSecuritySatisfaction',
                        'WorkLifeBalanceSatisfaction',
                        'GrowthOpportunitySatisfaction',
                        'WorkCultureSatisfaction',
                        'WorkDiversitySatisfaction',
                        'WorkloadSatisfaction']
```



We will drop any rows with NaN values (likely from respondents not in a coding role currently) with the `.dropna()` method.

```
satisfaction_df = survey_df[satisfaction_columns]
satisfaction_df = satisfaction_df.dropna(axis=0)
satisfaction_df
```

	EarningSatisfaction	BenefitsSatisfaction	JobSecuritySatisfaction	WorkLifeBalanceSatisfaction	GrowthOpportuni
0	Somewhat satisfied	Very satisfied	Somewhat satisfied	Somewhat dissatisfied	
1	Very dissatisfied	Very dissatisfied	Very dissatisfied	Somewhat satisfied	Somewh:
2	Not Applicable	Not Applicable	Not Applicable	Not Applicable	N
4	Somewhat satisfied	Somewhat dissatisfied	Somewhat dissatisfied	Somewhat satisfied	Somewh:
5	Somewhat satisfied	Somewhat dissatisfied	Somewhat dissatisfied	Very satisfied	Somewh:
...	...	...	...	...	
18119	Not Applicable	Not Applicable	Not Applicable	Not Applicable	N
18120	Very dissatisfied	Somewhat satisfied	Somewhat satisfied	Somewhat dissatisfied	Ve
18123	Somewhat dissatisfied	Very dissatisfied	Somewhat satisfied	Very satisfied	Ve
18124	Not Applicable	Somewhat satisfied	Somewhat satisfied	Somewhat satisfied	
18125	Somewhat dissatisfied	Not Applicable	Not Applicable	Very dissatisfied	N

11802 rows × 8 columns

```
satisfaction_df.EarningSatisfaction.unique()
```

```
array(['Somewhat satisfied', 'Very dissatisfied', 'Not Applicable',
      'Very satisfied', 'Somewhat dissatisfied', 'I do not know'],
      dtype=object)
```

We will replace each entry in our dataset with a score based on what the respondents answered. For example, if a respondent answered with 'Very Satisfied', we will score that a 1. If they answered with 'Somewhat dissatisfied', we will score that a -0.5. We will then sum up each column and create a list with the total score for each aspect.

```
scores = {'Very satisfied':1, 'Somewhat satisfied':0.5, 'Very dissatisfied':-1,
          'Somewhat dissatisfied':-0.5, 'Not Applicable':0, 'I do not know':0}
```

```
satisfaction_df = satisfaction_df.replace(scores)
satisfaction_df
```

	EarningSatisfaction	BenefitsSatisfaction	JobSecuritySatisfaction	WorkLifeBalanceSatisfaction	GrowthOpportuni
0	0.5	1.0	0.5	-0.5	
1	-1.0	-1.0	-1.0	0.5	
2	0.0	0.0	0.0	0.0	

	EarningSatisfaction	BenefitsSatisfaction	JobSecuritySatisfaction	WorkLifeBalanceSatisfaction	GrowthOpportuni
4	0.5	-0.5	-0.5	0.5	
5	0.5	-0.5	-0.5	1.0	
...	...	...	...	...	...
18119	0.0	0.0	0.0	0.0	
18120	-1.0	0.5	0.5	-0.5	
18123	-0.5	-1.0	0.5	1.0	
18124	0.0	0.5	0.5	0.5	
18125	-0.5	0.0	0.0	-1.0	

11802 rows × 8 columns

```
list(satisfaction_df.sum())
```

```
[-698.0, -103.0, 2005.5, 689.5, -1759.0, 405.0, 1292.5, 810.0]
```

```
results_list = list(satisfaction_df.sum())
```

```
print(results_list)
```

```
results_list.sort()
```

```
print(results_list)
```

```
[-698.0, -103.0, 2005.5, 689.5, -1759.0, 405.0, 1292.5, 810.0]
```

```
[-1759.0, -698.0, -103.0, 405.0, 689.5, 810.0, 1292.5, 2005.5]
```

We can now plot our results to see what respondents were most and least satisfied about in their jobs. We will use a Lollipop chart to visualise this. We will create some labels and use a for loop to set the color of our lines based on whether the score is positive or negative.

```
xlabel = "Very Dissatisfied"
ylabels = ['Growth Opportunity', 'Earnings',
           'Benefits', 'Work Culture',
           'Work-Life Balance', 'Workload',
           'Diversity', 'Job Security'] # need to order these in a specific way based on

line_dot_colors = [] # for our plot
for num in results_list:
    if num > 0:
        line_dot_colors.append('green')
    else:
        line_dot_colors.append('red')
```

```
fig, axes = plt.subplots()

axes.hlines(ylabels, xmin=0,
           xmax=results_list, colors=line_dot_colors)
plt.scatter(x=results_list, y=ylabels, s=75, color = line_dot_colors)
```

```
plt.title("How satisfied are you with different aspects of your job?")
plt.tick_params(labelbottom = False)
plt.ylabel(None)
plt.xlabel(xlabel);
```



The results indicate that job security is the aspect that coders are most satisfied with about their job, with diversity, workload, work-life balance and work culture also being positive points.

Growth opportunity is the aspect that the respondents were the most displeased about with their current jobs, with most also feeling that the earnings and benefits of their coding jobs could be improved upon.

## Q: Which regions are the most satisfied with their work conditions?

Finally, we will take our findings from the last question and go into further detail to find out how respondent's from different regions around the world feel about their current roles.

```
survey_df.Region.unique()
```

```
array(['Europe and Central Asia', 'Latin America and Caribbean',
      'North America', 'Sub-Saharan Africa', 'East Asia and Pacific',
      'South Asia', nan, 'Southeast Asia',
      'Middle East and North Africa'], dtype=object)
```

```
selected_columns = [ 'Region',
                    'EarningSatisfaction',
                    'BenefitsSatisfaction',
                    'JobSecuritySatisfaction',
                    'WorkLifeBalanceSatisfaction',
                    'GrowthOpportunitySatisfaction',
                    'WorkCultureSatisfaction',
                    'WorkDiversitySatisfaction',
                    'WorkloadSatisfaction' ]
```

```
regions_df = survey_df[selected_columns]
regions_df
```

	Region	EarningSatisfaction	BenefitsSatisfaction	JobSecuritySatisfaction	WorkLifeBalanceSatisfaction	Growth
0	Europe and Central Asia	Somewhat satisfied	Very satisfied	Somewhat satisfied	Somewhat dissatisfied	
1	Latin America and Caribbean	Very dissatisfied	Very dissatisfied	Very dissatisfied	Somewhat satisfied	
2	North America	Not Applicable	Not Applicable	Not Applicable	Not Applicable	
3	Sub-Saharan Africa	NaN	NaN	NaN	NaN	
4	East Asia and Pacific	Somewhat satisfied	Somewhat dissatisfied	Somewhat dissatisfied	Somewhat satisfied	
...	...	...	...	...	...	...
18121	Middle East and North Africa	NaN	NaN	NaN	NaN	
18122	North America	NaN	NaN	NaN	NaN	
18123	North America	Somewhat dissatisfied	Very dissatisfied	Somewhat satisfied	Very satisfied	
18124	Europe and Central Asia	Not Applicable	Somewhat satisfied	Somewhat satisfied	Somewhat satisfied	
18125	Europe and Central Asia	Somewhat dissatisfied	Not Applicable	Not Applicable	Very dissatisfied	

17938 rows × 9 columns

```
scores = {'Very satisfied':1, 'Somewhat satisfied':0.5, 'Very dissatisfied':-1,
          'Somewhat dissatisfied':-0.5, 'Not Applicable':0, 'I do not know':0}
regions_satisfaction_df = regions_df.replace(scores)
regions_satisfaction_df
```

	Region	EarningSatisfaction	BenefitsSatisfaction	JobSecuritySatisfaction	WorkLifeBalanceSatisfaction	Growth
0	Europe and Central Asia	0.5	1.0	0.5	-0.5	
1	Latin America and Caribbean	-1.0	-1.0	-1.0	0.5	

	Region	EarningSatisfaction	BenefitsSatisfaction	JobSecuritySatisfaction	WorkLifeBalanceSatisfaction	GrowthOpportunitySatisfaction
2	North America	0.0	0.0	0.0	0.0	0.0
3	Sub-Saharan Africa	NaN	NaN	NaN	NaN	NaN
4	East Asia and Pacific	0.5	-0.5	-0.5	0.5	0.5
...	...	...	...	...	...	...
18121	Middle East and North Africa	NaN	NaN	NaN	NaN	NaN
18122	North America	NaN	NaN	NaN	NaN	NaN
18123	North America	-0.5	-1.0	0.5	1.0	1.0
18124	Europe and Central Asia	0.0	0.5	0.5	0.5	0.5
18125	Europe and Central Asia	-0.5	0.0	0.0	-1.0	-1.0

17938 rows × 9 columns

```
regions_diff_df = regions_satisfaction_df.groupby('Region')[selected_columns[1:]].mean()
regions_diff_df = regions_diff_df.transpose()
regions_diff_df
```

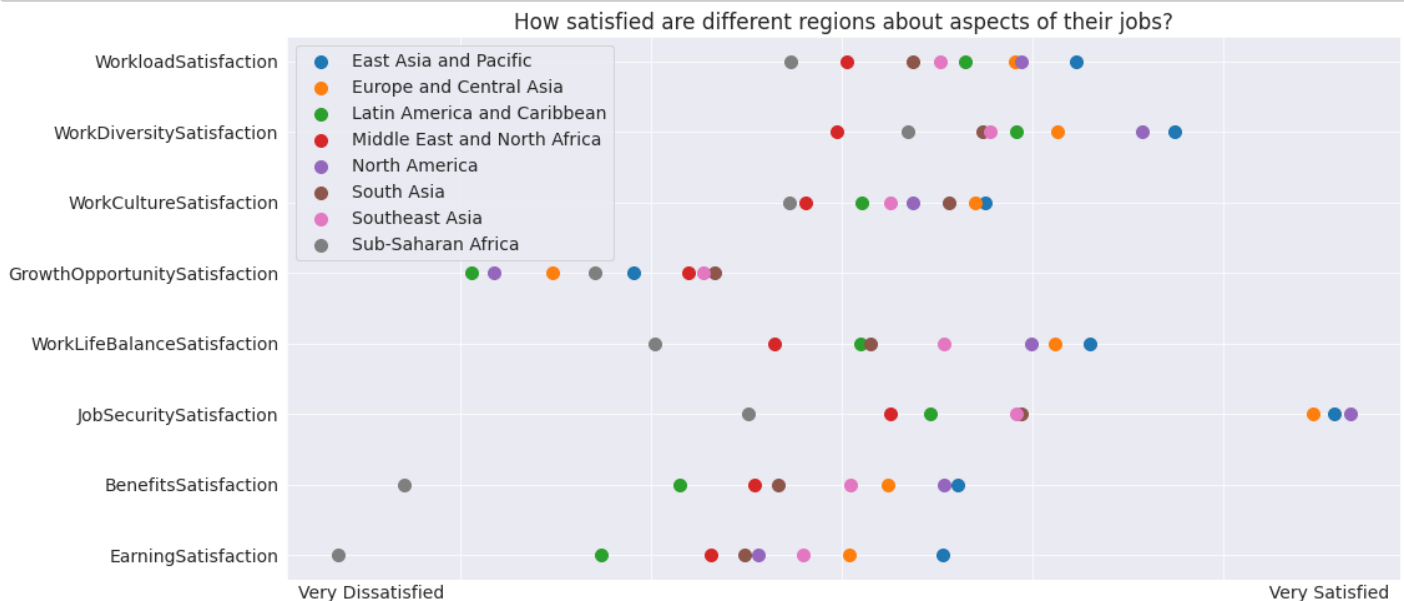
Region	East Asia and Pacific	Europe and Central Asia	Latin America and Caribbean	Middle East and North Africa	North America	South Asia	Southeast Asia	Sub-Saharan Africa
EarningSatisfaction	0.053218	0.003846	-0.125639	-0.068228	-0.043625	-0.050417	-0.019833	-0.2638
BenefitsSatisfaction	0.060794	0.024207	-0.084827	-0.045455	0.053732	-0.033102	0.004762	-0.2292
JobSecuritySatisfaction	0.258663	0.247148	0.046945	0.025988	0.266709	0.094470	0.091631	-0.0485
WorkLifeBalanceSatisfaction	0.130597	0.112129	0.010296	-0.034979	0.099724	0.015012	0.053968	-0.0979
GrowthOpportunitySatisfaction	-0.109181	-0.151201	-0.194140	-0.080208	-0.182400	-0.066512	-0.072487	-0.1293
WorkCultureSatisfaction	0.075682	0.070411	0.010925	-0.018595	0.037814	0.056327	0.025953	-0.0269
WorkDiversitySatisfaction	0.174564	0.113284	0.091495	-0.002070	0.157445	0.073930	0.078307	0.0350
WorkloadSatisfaction	0.122829	0.090835	0.064683	0.003093	0.094598	0.037568	0.051761	-0.0264

We will visualise our results using a scatterplot.

```
plt.figure(figsize=(16,8))

for col in regions_diff_df.columns:
    plt.scatter(regions_diff_df[col], regions_diff_df.index, s=100)
```

```
plt.title("How satisfied are different regions about aspects of their jobs?")
plt.xlabel("Very Dissatisfied
plt.tick_params(labelbottom = False)
plt.ylabel(None)
plt.legend(['East Asia and Pacific', 'Europe and Central Asia', 'Latin America and Caribbean',
            'North America', 'South Asia', 'Southeast Asia', 'Sub-Saharan Africa']);
```



Looking at the graph, East Asia and Pacific scored at the top of nearly all aspects, with job security being the aspect respondents were most satisfied with. Regarding job security, there appears to be a big gulf between respondents from East Asia and Pacific, Europe and Central Asia, and North America compared to the rest of the respondents.

Sub-Saharan Africa scored poorly in virtually all aspects, and dissatisfaction with benefits and earnings is particularly noteworthy.

Growth opportunity scored poorly in all regions, and work culture had the least variation compared to the rest, although satisfaction for this aspect was fairly average.

Let us save and upload our work to Jovian before continuing.

```
import jovian
```

```
jovian.commit()
```

```
[jovian] Updating notebook "adammasongh/2021-new-coder-survey" on https://jovian.com
[jovian] Committed successfully! https://jovian.com/adammasongh/2021-new-coder-survey
'https://jovian.com/adammasongh/2021-new-coder-survey'
```

## Inferences and Conclusion

We have gone through a lot of analysis and drawn many inferences from the dataset. Below is a summary of the main findings:

- A large number of the respondents stated their biggest reason for learning how to code is to start their career or to change their current careers. Roughly 12% of respondents said they learn how to code mainly as a hobby.
- The programming community is fairly skewed towards males and predominantly white nations (though this may be because this survey was conducted in English and therefore limited the number of people from different regions around the world. For example, coders from South America were not represented in this survey). There was a fairly large percentage of female respondents in this survey, and there was good representation of different ethnic backgrounds, highlighting that efforts to support underrepresented communities are having a positive effect.
- Most respondents held a degree of some kind, with computer science being the most common. However, a large percentage did not hold any type of university degree.
- A bachelors degree or high school diploma was the most common response to what the highest degree/qualification held by the respondent was. There was a good variety of responses highlighting that there is no right or wrong path to becoming a developer.
- Online resources and books were the most common methods new coders use to learn about coding, with over 90% and 40% respectively saying they use those methods.
- Men tend to spend more time learning how to code and have spent more months programming compared to women. There was no difference in age or money spent learning to code.
- There were some notable differences between different ethnic groups. White people tended to be older and have been programming for longer than other groups. Southeast Asians spent by far more money than any other group, almost 3 times as much as South Asians who spent the second most.
- Most developers are confident they will not lose their job within the next 12 months, and most are uncertain how easy it would be to find a new job if they lost their current job.
- Job security is the aspect coders were most satisfied with about their jobs, with diversity and workload also scoring highly.
- Growth opportunity is the aspect coders were least satisfied with about their jobs. Respondents believed earnings and benefits could also be improved.
- East Asia and Pacific scored the highest on nearly all aspects of job satisfaction. Sub-Saharan Africa scored poorly on every aspect, with benefits and earnings scoring particularly low.
- Job security was scored highly by respondents from East Asia and Pacific, North America, and Europe and Central Asia. There a notable gulf between these regions and other regions regarding this aspect.

```
import jovian
```

```
jovian.commit()
```

```
[jovian] Updating notebook "adammasongh/2021-new-coder-survey" on https://jovian.com  
[jovian] Committed successfully! https://jovian.com/adammasongh/2021-new-coder-survey  
'https://jovian.com/adammasongh/2021-new-coder-survey'
```

## References and Future Work

The original dataset obtained from Kaggle contains more columns of data which weren't covered in this project but could be analysed in the future to gather more insight into the respondents.

It would be useful if future surveys asked questions about programming languages so we could get more information relating to coder's preferred languages, the most popular languages, the fastest growing languages etc. This would provide new coders valuable knowledge about the best programming languages to use in their future careers. The annual Stack Overflow Developer Survey asks respondents questions relating to the programming languages they use, so analysing that data could provide us more insight into this aspect of coding.

It would also be good to learn more about people currently in developer jobs so we could get a better idea about their daily work routine, hours they work, salary etc so we could understand more about the life of current developers. As we found out from the analysis, earnings and benefits were seen as areas that could be improved, so knowing more about these aspects could help us find out the specifics as to why these are rating so poorly by developers.

I have added the links to the following resources that were used to carry out this project:

- Dataset used in project - <https://www.kaggle.com/datasets/fccuser/2021-new-coder-survey>
- 30 Examples to Master Python - <https://towardsdatascience.com/30-examples-to-master-pandas-f8a2da751fa4>
- Stack Overflow website - <https://stackoverflow.com/>
- Seaborn user guide & tutorial - <https://seaborn.pydata.org/tutorial.html>
- Pandas user guide: [https://pandas.pydata.org/docs/user\\_guide/index.html](https://pandas.pydata.org/docs/user_guide/index.html)
- Create a Dataframe from List - <https://datagy.io/pandas-dataframe-from-list/>
- Renaming columns - <https://www.geeksforgeeks.org/how-to-rename-columns-in-pandas-dataframe/>
- Top 50 Matplotlib Visualizations - <https://www.machinelearningplus.com/plots/top-50-matplotlib-visualizations-the-master-plots-python/>
- Grouped Bar Charts with Labels in Matplotlib - <https://www.pythoncharts.com/matplotlib/grouped-bar-charts-matplotlib/>
- Joining two dataframes together vertically: <https://stackoverflow.com/questions/41181779/merging-2-dataframes-vertically>
- How to split male and female respondents in a histogram: <https://stackoverflow.com/questions/51058053/how-to-plot-a-histogram-of-one-column-colored-by-another-in-python>

```
import jovian
```

```
jovian.commit()
```