In [ ]:

```
Answer 6
```

In [8]:

```
import selenium
import pandas as pd
import time
from bs4 import BeautifulSoup
from selenium import webdriver
from selenium.common.exceptions import StaleElementReferenceException,
NoSuchElementException
import requests
from selenium.webdriver.common.by import By
```

In [9]:

```
driver=webdriver.Chrome(r"C:\Users\zarna\Downloads\chromedriver_win32
(1)\chromedriver.exe")
```

```
C:\Users\zarna\AppData\Local\Temp\ipykernel_169148\1240552591.py:1:
DeprecationWarning: executable_path has been deprecated, please pass in a
Service object
  driver=webdriver.Chrome(r"C:\Users\zarna\Downloads\chromedriver_win32
(1)\chromedriver.exe")
```

In [11]:

```
driver.get("https://www.theguardian.com/news/datablog/2012/aug/09/best-
selling-books-all-time-fifty-shades-grey-compare")
```

In [12]:

```
Book_name=[]
Author_name=[]
Volumes_sold=[]
Publisher=[]
Genre=[]
```

In [13]:

```
data=[]

try:
    data_tags=driver.find_elements(By.XPATH,'//td[@class="left"]')
    for i in data_tags:
        data.append(i.text)
except NoSuchElementException:
    data.append('-')
```

In [14]:

```
data
```

Out[14]:

```
['1',
 'Da Vinci Code,The',
 'Brown, Dan',
 '5,094,805',
 'Transworld',
 '2',
 'Harry Potter and the Deathly Hallows',
 'Rowling, J.K.',
 '4,475,152',
```

```
'Bloomsbury',
'3',
"Harry Potter and the Philosopher's Stone",
'Rowling, J.K.',
'4,200,654',
'Bloomsbury',
'4',
'Harry Potter and the Order of the Phoenix',
'Rowling, J.K.',
'4,179,479',
'Bloomsbury',
'5',
'Fifty Shades of Grey',
'James, E. L.',
'3,758,936',
'Random House',
'6',
'Harry Potter and the Goblet of Fire',
'Rowling, J.K.',
'3,583,215',
'Bloomsbury',
'7',
'Harry Potter and the Chamber of Secrets',
'Rowling, J.K.',
'3,484,047',
'Bloomsbury',
'8',
'Harry Potter and the Prisoner of Azkaban',
'Rowling, J.K.',
'3,377,906',
'Bloomsbury',
'9',
'Angels and Demons',
'Brown, Dan',
'3,193,946',
'Transworld',
'10',
"Harry Potter and the Half-blood Prince:Children's Edition",
'Rowling, J.K.',
'2,950,264',
'Bloomsbury',
'11',
'Fifty Shades Darker',
'James, E. L.',
'2,479,784',
'Random House',
'12',
'Twilight',
'Meyer, Stephenie',
```

'2,315,405',
'Little, Brown Book',
'13',
'Girl with the Dragon Tattoo,The:Millennium Trilogy',
'Larsson, Stieg',
'2,233,570',
'Quercus',
'14',
'Fifty Shades Freed',
'James, E. L.',
'2,193,928',
'Random House',
'15',
'Lost Symbol,The',
'Brown, Dan',
'2,183,031',
'Transworld',
'16',
'New Moon',
'Meyer, Stephenie',
'2,152,737',
'Little, Brown Book',
'17',
'Deception Point',
'Brown, Dan',
'2,062,145',
'Transworld',
'18',
'Eclipse',
'Meyer, Stephenie',
'2,052,876',
'Little, Brown Book',
'19',
'Lovely Bones,The',
'Sebold, Alice',
'2,005,598',
'Pan Macmillan',
'20',
'Curious Incident of the Dog in the Night-time,The',
'Haddon, Mark',
'1,979,552',
'Random House',
'21',
'Digital Fortress',
'Brown, Dan',
'1,928,900',
'Transworld',
'22',
'Short History of Nearly Everything,A',

'Bryson, Bill',
'1,852,919',
'Transworld',
'23',
'Girl Who Played with Fire,The:Millennium Trilogy',
'Larsson, Stieg',
'1,814,784',
'Quercus',
'24',
'Breaking Dawn',
'Meyer, Stephenie',
'1,787,118',
'Little, Brown Book',
'25',
'Very Hungry Caterpillar,The:The Very Hungry Caterpillar',
'Carle, Eric',
'1,783,535',
'Penguin',
'26',
'Gruffalo,The',
'Donaldson, Julia',
'1,781,269',
'Pan Macmillan',
'27',
"Jamie's 30-Minute Meals",
'Oliver, Jamie',
'1,743,266',
'Penguin',
'28',
'Kite Runner,The',
'Hosseini, Khaled',
'1,629,119',
'Bloomsbury',
'29',
'One Day',
'Nicholls, David',
'1,616,068',
'Hodder & Stoughton',
'30',
'Thousand Splendid Suns,A',
'Hosseini, Khaled',
'1,583,992',
'Bloomsbury',
'31',
"Girl Who Kicked the Hornets' Nest,The:Millennium Trilogy",
'Larsson, Stieg',
'1,555,135',
'Quercus',
'32',

"Time Traveler's Wife,The",
'Niffenegger, Audrey',
'1,546,886',
'Random House',
'33',
'Atonement',
'McEwan, Ian',
'1,539,428',
'Random House',
'34',
"Bridget Jones's Diary:A Novel",
'Fielding, Helen',
'1,508,205',
'Pan Macmillan',
'35',
'World According to Clarkson,The',
'Clarkson, Jeremy',
'1,489,403',
'Penguin',
'36',
"Captain Corelli's Mandolin",
'Bernieres, Louis de',
'1,352,318',
'Random House',
'37',
'Sound of Laughter,The',
'Kay, Peter',
'1,310,207',
'Random House',
'38',
'Life of Pi',
'Martel, Yann',
'1,310,176',
'Canongate',
'39',
'Billy Connolly',
'Stephenson, Pamela',
'1,231,957',
'HarperCollins',
'40',
'Child Called It,A',
'Pelzer, Dave',
'1,217,712',
'Orion',
'41',
"Gruffalo's Child,The",
'Donaldson, Julia',
'1,208,711',
'Pan Macmillan',

'42',
"Angela's Ashes:A Memoir of a Childhood",
'McCourt, Frank',
'1,204,058',
'HarperCollins',
'43',
'Birdsong',
'Faulks, Sebastian',
'1,184,967',
'Random House',
'44',
'Northern Lights:His Dark Materials S.',
'Pullman, Philip',
'1,181,503',
'Scholastic Ltd.',
'45',
'Labyrinth',
'Mosse, Kate',
'1,181,093',
'Orion',
'46',
'Harry Potter and the Half-blood Prince',
'Rowling, J.K.',
'1,153,181',
'Bloomsbury',
'47',
'Help,The',
'Stockett, Kathryn',
'1,132,336',
'Penguin',
'48',
'Man and Boy',
'Parsons, Tony',
'1,130,802',
'HarperCollins',
'49',
'Memoirs of a Geisha',
'Golden, Arthur',
'1,126,337',
'Random House',
'50',
"No.1 Ladies' Detective Agency,The:No.1 Ladies' Detective Agency S.",
'McCall Smith, Alexander',
'1,115,549',
'Little, Brown Book',
'51',
'Island,The',
'Hislop, Victoria',
'1,108,328',

'Headline',
'52',
'PS, I Love You',
'Ahern, Cecelia',
'1,107,379',
'HarperCollins',
'53',
'You are What You Eat:The Plan That Will Change Your Life',
'McKeith, Gillian',
'1,104,403',
'Penguin',
'54',
'Shadow of the Wind,The',
'Zafon, Carlos Ruiz',
'1,092,349',
'Orion',
'55',
'Tales of Beedle the Bard,The',
'Rowling, J.K.',
'1,090,847',
'Bloomsbury',
'56',
'Broker,The',
'Grisham, John',
'1,087,262',
'Random House',
'57',
"Dr. Atkins' New Diet Revolution:The No-hunger, Luxurious Weight Loss P",
'Atkins, Robert C.',
'1,054,196',
'Random House',
'58',
'Subtle Knife,The:His Dark Materials S.',
'Pullman, Philip',
'1,037,160',
'Scholastic Ltd.',
'59',
'Eats, Shoots and Leaves:The Zero Tolerance Approach to Punctuation',
'Truss, Lynne',
'1,023,688',
'Profile Books Group',
'60',
"Delia's How to Cook:(Bk.1)",
'Smith, Delia',
'1,015,956',
'Random House',
'61',
'Chocolat',
'Harris, Joanne',

'1,009,873',
'Transworld',
'62',
'Boy in the Striped Pyjamas,The',
'Boyne, John',
'1,004,414',
'Random House Childrens Books G',
'63',
"My Sister's Keeper",
'Picoult, Jodi',
'1,003,780',
'Hodder & Stoughton',
'64',
'Amber Spyglass,The:His Dark Materials S.',
'Pullman, Philip',
'1,002,314',
'Scholastic Ltd.',
'65',
'To Kill a Mockingbird',
'Lee, Harper',
'998,213',
'Random House',
'66',
'Men are from Mars, Women are from Venus:A Practical Guide for Improvin',
'Gray, John',
'992,846',
'HarperCollins',
'67',
'Dear Fatty',
'French, Dawn',
'986,753',
'Random House',
'68',
'Short History of Tractors in Ukrainian,A',
'Lewycka, Marina',
'986,115',
'Penguin',
'69',
'Hannibal',
'Harris, Thomas',
'970,509',
'Random House',
'70',
'Lord of the Rings,The',
'Tolkien, J. R. R.',
'967,466',
'HarperCollins',
'71',
'Stupid White Men:...and Other Sorry Excuses for the State of the Natio',

'Moore, Michael',
'963,353',
'Penguin',
'72',
'Interpretation of Murder,The',
'Rubenfeld, Jed',
'962,515',
'Headline',
'73',
'Sharon Osbourne Extreme:My Autobiography',
'Osbourne, Sharon',
'959,496',
'Little, Brown Book',
'74',
'Alchemist,The:A Fable About Following Your Dream',
'Coelho, Paulo',
'956,114',
'HarperCollins',
'75',
"At My Mother's Knee ...:and Other Low Joints",
"O'Grady, Paul",
'945,640',
'Transworld',
'76',
'Notes from a Small Island',
'Bryson, Bill',
'931,312',
'Transworld',
'77',
'Return of the Naked Chef,The',
'Oliver, Jamie',
'925,425',
'Penguin',
'78',
'Bridget Jones: The Edge of Reason',
'Fielding, Helen',
'924,695',
'Pan Macmillan',
'79',
"Jamie's Italy",
'Oliver, Jamie',
'906,968',
'Penguin',
'80',
'I Can Make You Thin',
'McKenna, Paul',
'905,086',
'Transworld',
'81',

'Down Under',
'Bryson, Bill',
'890,847',
'Transworld',
'82',
'Summons,The',
'Grisham, John',
'869,671',
'Random House',
'83',
'Small Island',
'Levy, Andrea',
'869,659',
'Headline',
'84',
'Nigella Express',
'Lawson, Nigella',
'862,602',
'Random House',
'85',
'Brick Lane',
'Ali, Monica',
'856,540',
'Transworld',
'86',
"Memory Keeper's Daughter,The",
'Edwards, Kim',
'845,858',
'Penguin',
'87',
'Room on the Broom',
'Donaldson, Julia',
'842,535',
'Pan Macmillan',
'88',
'About a Boy',
'Hornby, Nick',
'828,215',
'Penguin',
'89',
'My Booky Wook',
'Brand, Russell',
'820,563',
'Hodder & Stoughton',
'90',
'God Delusion,The',
'Dawkins, Richard',
'816,907',
'Transworld',

'91',
'"Beano" Annual,The',
'0',
'816,585',
'D.C. Thomson',
'92',
'White Teeth',
'Smith, Zadie',
'815,586',
'Penguin',
'93',
'House at Riverton,The',
'Morton, Kate',
'814,370',
'Pan Macmillan',
'94',
'Book Thief,The',
'Zusak, Markus',
'809,641',
'Transworld',
'95',
'Nights of Rain and Stars',
'Binchy, Maeve',
'808,900',
'Orion',
'96',
'Ghost,The',
'Harris, Robert',
'807,311',
'Random House',
'97',
'Happy Days with the Naked Chef',
'Oliver, Jamie',
'794,201',
'Penguin',
'98',
'Hunger Games,The:Hunger Games Trilogy',
'Collins, Suzanne',
'792,187',
'Scholastic Ltd.',
'99',
"Lost Boy,The:A Foster Child's Search for the Love of a Family",
'Pelzer, Dave',
'791,507',
'Orion',
'100',
"Jamie's Ministry of Food:Anyone Can Learn to Cook in 24 Hours",
'Oliver, Jamie',
'791,095',

```
 'Penguin']
```

In [15]:

```
Book_name=data[1::5]
Book_name
```

Out[15]:

```
['Da Vinci Code,The',
 'Harry Potter and the Deathly Hallows',
 "Harry Potter and the Philosopher's Stone",
 'Harry Potter and the Order of the Phoenix',
 'Fifty Shades of Grey',
 'Harry Potter and the Goblet of Fire',
 'Harry Potter and the Chamber of Secrets',
 'Harry Potter and the Prisoner of Azkaban',
 'Angels and Demons',
 "Harry Potter and the Half-blood Prince:Children's Edition",
 'Fifty Shades Darker',
 'Twilight',
 'Girl with the Dragon Tattoo,The:Millennium Trilogy',
 'Fifty Shades Freed',
 'Lost Symbol,The',
 'New Moon',
 'Deception Point',
 'Eclipse',
 'Lovely Bones,The',
 'Curious Incident of the Dog in the Night-time,The',
 'Digital Fortress',
 'Short History of Nearly Everything,A',
 'Girl Who Played with Fire,The:Millennium Trilogy',
 'Breaking Dawn',
 'Very Hungry Caterpillar,The:The Very Hungry Caterpillar',
 'Gruffalo,The',
 "Jamie's 30-Minute Meals",
 'Kite Runner,The',
 'One Day',
 'Thousand Splendid Suns,A',
 "Girl Who Kicked the Hornets' Nest,The:Millennium Trilogy",
 "Time Traveler's Wife,The",
 'Atonement',
 "Bridget Jones's Diary:A Novel",
 'World According to Clarkson,The',
 "Captain Corelli's Mandolin",
 'Sound of Laughter,The',
 'Life of Pi',
 'Billy Connolly',
 'Child Called It,A',
 "Gruffalo's Child,The",
 "Angela's Ashes:A Memoir of a Childhood",
 'Birdsong',
 'Northern Lights:His Dark Materials S.',
```

'Labyrinth',
'Harry Potter and the Half-blood Prince',
'Help,The',
'Man and Boy',
'Memoirs of a Geisha',
"No.1 Ladies' Detective Agency,The:No.1 Ladies' Detective Agency S.",
'Island,The',
'PS, I Love You',
'You are What You Eat:The Plan That Will Change Your Life',
'Shadow of the Wind,The',
'Tales of Beedle the Bard,The',
'Broker,The',
"Dr. Atkins' New Diet Revolution:The No-hunger, Luxurious Weight Loss P",
'Subtle Knife,The:His Dark Materials S.',
'Eats, Shoots and Leaves:The Zero Tolerance Approach to Punctuation',
"Delia's How to Cook:(Bk.1)",
'Chocolat',
'Boy in the Striped Pyjamas,The',
"My Sister's Keeper",
'Amber Spyglass,The:His Dark Materials S.',
'To Kill a Mockingbird',
'Men are from Mars, Women are from Venus:A Practical Guide for Improvin',
'Dear Fatty',
'Short History of Tractors in Ukrainian,A',
'Hannibal',
'Lord of the Rings,The',
'Stupid White Men:...and Other Sorry Excuses for the State of the Natio',
'Interpretation of Murder,The',
'Sharon Osbourne Extreme:My Autobiography',
'Alchemist,The:A Fable About Following Your Dream',
"At My Mother's Knee ...:and Other Low Joints",
'Notes from a Small Island',
'Return of the Naked Chef,The',
'Bridget Jones: The Edge of Reason',
"Jamie's Italy",
'I Can Make You Thin',
'Down Under',
'Summons,The',
'Small Island',
'Nigella Express',
'Brick Lane',
"Memory Keeper's Daughter,The",
'Room on the Broom',
'About a Boy',
'My Booky Wook',
'God Delusion,The',
'"Beano" Annual,The',
'White Teeth',
'House at Riverton,The',

```
 'Book Thief,The',
 'Nights of Rain and Stars',
 'Ghost,The',
 'Happy Days with the Naked Chef',
 'Hunger Games,The:Hunger Games Trilogy',
 "Lost Boy,The:A Foster Child's Search for the Love of a Family",
 "Jamie's Ministry of Food:Anyone Can Learn to Cook in 24 Hours"]
```

In [16]:

```
Author_name=data[2::5]
Volumes_sold=data[3::5]
Publisher=data[4::5]
```

In [17]:

```
try:
    genre_tags=driver.find_elements(By.XPATH,'//td[@class="last left"]')
    for i in genre_tags:
        Genre.append(i.text)
except NoSuchElementException:
    Genre.append("-")
```

In [18]:

```
Genre
```

Out[18]:

```
['Crime, Thriller & Adventure',
 "Children's Fiction",
 "Children's Fiction",
 "Children's Fiction",
 'Romance & Sagas',
 "Children's Fiction",
 "Children's Fiction",
 "Children's Fiction",
 'Crime, Thriller & Adventure',
 "Children's Fiction",
 'Romance & Sagas',
 'Young Adult Fiction',
 'Crime, Thriller & Adventure',
 'Romance & Sagas',
 'Crime, Thriller & Adventure',
 'Young Adult Fiction',
 'Crime, Thriller & Adventure',
 'Young Adult Fiction',
 'General & Literary Fiction',
 'General & Literary Fiction',
 'Crime, Thriller & Adventure',
 'Popular Science',
 'Crime, Thriller & Adventure',
 'Young Adult Fiction',
 'Picture Books',
 'Picture Books',
 'Food & Drink: General',
 'General & Literary Fiction',
 'General & Literary Fiction',
```

```
'General & Literary Fiction',
'Crime, Thriller & Adventure',
'General & Literary Fiction',
'General & Literary Fiction',
'General & Literary Fiction',
'Humour: Collections & General',
'General & Literary Fiction',
'Autobiography: General',
'General & Literary Fiction',
'Biography: The Arts',
'Autobiography: General',
'Picture Books',
'Autobiography: General',
'General & Literary Fiction',
'Young Adult Fiction',
'General & Literary Fiction',
'Science Fiction & Fantasy',
'General & Literary Fiction',
'General & Literary Fiction',
'General & Literary Fiction',
'Crime, Thriller & Adventure',
'General & Literary Fiction',
'General & Literary Fiction',
'Fitness & Diet',
'General & Literary Fiction',
"Children's Fiction",
'Crime, Thriller & Adventure',
'Fitness & Diet',
'Young Adult Fiction',
'Usage & Writing Guides',
'Food & Drink: General',
'General & Literary Fiction',
'Young Adult Fiction',
'General & Literary Fiction',
'Young Adult Fiction',
'General & Literary Fiction',
'Popular Culture & Media: General Interest',
'Autobiography: The Arts',
'General & Literary Fiction',
'Crime, Thriller & Adventure',
'Science Fiction & Fantasy',
'Current Affairs & Issues',
'Crime, Thriller & Adventure',
'Autobiography: The Arts',
'General & Literary Fiction',
'Autobiography: The Arts',
'Travel Writing',
'Food & Drink: General',
'General & Literary Fiction',
```

```
 'National & Regional Cuisine',
 'Fitness & Diet',
 'Travel Writing',
 'Crime, Thriller & Adventure',
 'General & Literary Fiction',
 'Food & Drink: General',
 'General & Literary Fiction',
 'General & Literary Fiction',
 'Picture Books',
 'General & Literary Fiction',
 'Autobiography: The Arts',
 'Popular Science',
 "Children's Annuals",
 'General & Literary Fiction',
 'General & Literary Fiction',
 'General & Literary Fiction',
 'General & Literary Fiction',
 'General & Literary Fiction',
 'Food & Drink: General',
 'Young Adult Fiction',
 'Biography: General',
 'Food & Drink: General']
```

In [19]:

```
print(len(Book_name),len(Author_name),len(Volumes_sold),len(Publisher),len(Genre))
```

```
100 100 100 100 100
```

In [20]:

```
import pandas as pd
db=pd.DataFrame({"Book name":Book_name,"Author name":Author_name,"Volumes Sold":Volumes_sold,"Publisher":Publisher,"Genre":Genre})
db
```

Out[20]:

	Book name	Author name	Volumes Sold	Publisher	Genre
0	Da Vinci Code,The	Brown, Dan	5,094,805	Transworld	Crime, Thriller & Adventure
1	Harry Potter and the Deathly Hallows	Rowling, J.K.	4,475,152	Bloomsbury	Children's Fiction
2	Harry Potter and the Philosopher's Stone	Rowling, J.K.	4,200,654	Bloomsbury	Children's Fiction
3	Harry Potter and the Order of the Phoenix	Rowling, J.K.	4,179,479	Bloomsbury	Children's Fiction
4	Fifty Shades of Grey	James, E. L.	3,758,936	Random House	Romance & Sagas
...	...	...	...	...	...
95	Ghost,The	Harris, Robert	807,311	Random House	General & Literary Fiction

	Book name	Author name	Volumes Sold	Publisher	Genre
96	Happy Days with the Naked Chef	Oliver, Jamie	794,201	Penguin	Food & Drink: General
97	Hunger Games,The:Hunger Games Trilogy	Collins, Suzanne	792,187	Scholastic Ltd.	Young Adult Fiction
98	Lost Boy,The:A Foster Child's Search for the L...	Pelzer, Dave	791,507	Orion	Biography: General
99	Jamie's Ministry of Food:Anyone Can Learn to C...	Oliver, Jamie	791,095	Penguin	Food & Drink: General

100 rows × 5 columns

In [ ]:

```
Answer 7
```

In [21]:

```
import selenium
import pandas as pd
import time
from bs4 import BeautifulSoup
from selenium import webdriver
from selenium.common.exceptions import StaleElementReferenceException,
NoSuchElementException
import requests
from selenium.webdriver.common.by import By
```

In [22]:

```
driver=webdriver.Chrome(r"C:\Users\zarna\Downloads\chromedriver_win32
(1)\chromedriver.exe")
C:\Users\zarna\AppData\Local\Temp\ipykernel_169148\1240552591.py:1:
DeprecationWarning: executable_path has been deprecated, please pass in a
Service object
  driver=webdriver.Chrome(r"C:\Users\zarna\Downloads\chromedriver_win32
(1)\chromedriver.exe")
```

In [23]:

```
driver.get("https://www.imdb.com/list/ls095964455/")
```

In [24]:

```
Name=[]
Year_span=[]
genre=[]
Run_time=[]
Ratings=[]
Votes=[]
```

In [25]:

```
data=[]

try:
    data_tags=driver.find_elements(By.XPATH,'//h3[@class="lister-item-
header"]')
    for i in data_tags:
        data.append(i.text)
except NoSuchElementException:
```

```
    data.append("-")
```

```
data
```

Out[25]:

```
['1. Game of Thrones (2011–2019)',
 '2. Stranger Things (2016–2024)',
 '3. The Walking Dead (2010–2022)',
 '4. 13 Reasons Why (2017–2020)',
 '5. The 100 (2014–2020)',
 '6. Orange Is the New Black (2013–2019)',
 '7. Riverdale (2017–2023)',
 "8. Grey's Anatomy (2005– )",
 '9. The Flash (2014–2023)',
 '10. Arrow (2012–2020)',
 '11. Money Heist (2017–2021)',
 '12. The Big Bang Theory (2007–2019)',
 '13. Black Mirror (2011– )',
 '14. Sherlock (2010–2017)',
 '15. Vikings (2013–2020)',
 '16. Pretty Little Liars (2010–2017)',
 '17. The Vampire Diaries (2009–2017)',
 '18. American Horror Story (2011– )',
 '19. Breaking Bad (2008–2013)',
 '20. Lucifer (2016–2021)',
 '21. Supernatural (2005–2020)',
 '22. Prison Break (2005–2017)',
 '23. How to Get Away with Murder (2014–2020)',
 '24. Teen Wolf (2011–2017)',
 '25. The Simpsons (1989– )',
 '26. Once Upon a Time (2011–2018)',
 '27. Narcos (2015–2017)',
 '28. Daredevil (2015–2018)',
 '29. Friends (1994–2004)',
 '30. How I Met Your Mother (2005–2014)',
 '31. Suits (2011–2019)',
 '32. Mr. Robot (2015–2019)',
 '33. The Originals (2013–2018)',
 '34. Supergirl (2015–2021)',
 '35. Gossip Girl (2007–2012)',
 '36. Sense8 (2015–2018)',
 '37. Gotham (2014–2019)',
 '38. Westworld (2016–2022)',
 '39. Jessica Jones (2015–2019)',
 '40. Modern Family (2009–2020)',
 '41. Rick and Morty (2013– )',
 '42. Shadowhunters (2016–2019)',
 '43. The End of the F***ing World (2017–2019)',
 '44. House of Cards (2013–2018)',
 '45. Dark (2017–2020)',
```

'46. Elite (2018-)',
'47. Sex Education (2019-)',
'48. Shameless (2011-2021)',
'49. New Girl (2011-2018)',
'50. Agents of S.H.I.E.L.D. (2013-2020)',
'51. You (2018-2024)',
'52. Dexter (2006-2013)',
'53. Fear the Walking Dead (2015-2023)',
'54. Family Guy (1999-)',
'55. The Blacklist (2013-2023)',
'56. Lost (2004-2010)',
'57. Peaky Blinders (2013-2022)',
'58. House (2004-2012)',
'59. Quantico (2015-2018)',
'60. Orphan Black (2013-2017)',
'61. Homeland (2011-2020)',
'62. Blindspot (2015-2020)',
"63. DC's Legends of Tomorrow (2016-2022)",
"64. The Handmaid's Tale (2017-)",
'65. Chilling Adventures of Sabrina (2018-2020)',
'66. The Good Doctor (2017-)',
'67. Jane the Virgin (2014-2019)',
'68. Glee (2009-2015)',
'69. South Park (1997-)',
'70. Brooklyn Nine-Nine (2013-2021)',
'71. Under the Dome (2013-2015)',
'72. The Umbrella Academy (2019-2023)',
'73. True Detective (2014-)',
'74. The OA (2016-2019)',
'75. Desperate Housewives (2004-)',
'76. Better Call Saul (2015-2022)',
'77. Bates Motel (2013-2017)',
'78. The Punisher (2017-2019)',
'79. Atypical (2017-2021)',
'80. Dynasty (2017-2022)',
'81. This Is Us (2016-2022)',
'82. The Good Place (2016-2020)',
'83. Iron Fist (2017-2018)',
'84. The Rain (2018-2020)',
'85. Mindhunter (2017-2019)',
'86. Revenge (2011-2015)',
'87. Luke Cage (2016-2018)',
'88. Scandal (2012-2018)',
'89. The Defenders (2017)',
'90. Big Little Lies (2017-2019)',
'91. Insatiable (2018-2019)',
'92. The Mentalist (2008-2015)',
'93. The Crown (2016-2023)',
'94. Chernobyl (2019)',

```
 '95. iZombie (2015–2019)',
 '96. Reign (2013–2017)',
 '97. A Series of Unfortunate Events (2017–2019)',
 '98. Criminal Minds (2005– )',
 '99. Scream: The TV Series (2015–2019)',
 '100. The Haunting of Hill House (2018)']
```

In [26]:

```
try:
    year_tags=driver.find_elements(By.XPATH,'//span[@class="lister-item-
year text-muted unbold"]')
    for i in year_tags:
        Year_span.append(i.text)
except NoSuchElementException:
    Year_span.append("-")

Year_span
```

Out[26]:

```
['(2011–2019)',
 '(2016–2024)',
 '(2010–2022)',
 '(2017–2020)',
 '(2014–2020)',
 '(2013–2019)',
 '(2017–2023)',
 '(2005– )',
 '(2014–2023)',
 '(2012–2020)',
 '(2017–2021)',
 '(2007–2019)',
 '(2011– )',
 '(2010–2017)',
 '(2013–2020)',
 '(2010–2017)',
 '(2009–2017)',
 '(2011– )',
 '(2008–2013)',
 '(2016–2021)',
 '(2005–2020)',
 '(2005–2017)',
 '(2014–2020)',
 '(2011–2017)',
 '(1989– )',
 '(2011–2018)',
 '(2015–2017)',
 '(2015–2018)',
 '(1994–2004)',
 '(2005–2014)',
 '(2011–2019)',
 '(2015–2019)',
```

'(2013-2018)',
'(2015-2021)',
'(2007-2012)',
'(2015-2018)',
'(2014-2019)',
'(2016-2022)',
'(2015-2019)',
'(2009-2020)',
'(2013-)',
'(2016-2019)',
'(2017-2019)',
'(2013-2018)',
'(2017-2020)',
'(2018-)',
'(2019-)',
'(2011-2021)',
'(2011-2018)',
'(2013-2020)',
'(2018-2024)',
'(2006-2013)',
'(2015-2023)',
'(1999-)',
'(2013-2023)',
'(2004-2010)',
'(2013-2022)',
'(2004-2012)',
'(2015-2018)',
'(2013-2017)',
'(2011-2020)',
'(2015-2020)',
'(2016-2022)',
'(2017-)',
'(2018-2020)',
'(2017-)',
'(2014-2019)',
'(2009-2015)',
'(1997-)',
'(2013-2021)',
'(2013-2015)',
'(2019-2023)',
'(2014-)',
'(2016-2019)',
'(2004-)',
'(2015-2022)',
'(2013-2017)',
'(2017-2019)',
'(2017-2021)',
'(2017-2022)',
'(2016-2022)',

```
 '(2016–2020)',
 '(2017–2018)',
 '(2018–2020)',
 '(2017–2019)',
 '(2011–2015)',
 '(2016–2018)',
 '(2012–2018)',
 '(2017)',
 '(2017–2019)',
 '(2018–2019)',
 '(2008–2015)',
 '(2016–2023)',
 '(2019)',
 '(2015–2019)',
 '(2013–2017)',
 '(2017–2019)',
 '(2005– )',
 '(2015–2019)',
 '(2018)']
```

In [27]:

```
try:
    genre_tags=driver.find_elements(By.XPATH,'//span[@class="genre"]')
    for i in genre_tags:
        genre.append(i.text)
except NoSuchElementException:
    genre.append("-")

genre
```

Out[27]:

```
['Action, Adventure, Drama',
 'Drama, Fantasy, Horror',
 'Drama, Horror, Thriller',
 'Drama, Mystery, Thriller',
 'Drama, Mystery, Sci-Fi',
 'Comedy, Crime, Drama',
 'Crime, Drama, Mystery',
 'Drama, Romance',
 'Action, Adventure, Drama',
 'Action, Adventure, Crime',
 'Action, Crime, Drama',
 'Comedy, Romance',
 'Drama, Mystery, Sci-Fi',
 'Crime, Drama, Mystery',
 'Action, Adventure, Drama',
 'Drama, Mystery, Romance',
 'Drama, Fantasy, Horror',
 'Drama, Horror, Sci-Fi',
 'Crime, Drama, Thriller',
 'Crime, Drama, Fantasy',
```

```
'Drama, Fantasy, Horror',
'Action, Crime, Drama',
'Crime, Drama, Mystery',
'Action, Drama, Fantasy',
'Animation, Comedy',
'Adventure, Fantasy, Romance',
'Biography, Crime, Drama',
'Action, Crime, Drama',
'Comedy, Romance',
'Comedy, Drama, Romance',
'Comedy, Drama',
'Crime, Drama, Thriller',
'Drama, Fantasy, Horror',
'Action, Adventure, Drama',
'Drama, Romance',
'Drama, Mystery, Sci-Fi',
'Action, Crime, Drama',
'Drama, Mystery, Sci-Fi',
'Action, Crime, Drama',
'Comedy, Drama, Romance',
'Animation, Adventure, Comedy',
'Action, Drama, Fantasy',
'Adventure, Comedy, Crime',
'Drama',
'Crime, Drama, Mystery',
'Crime, Drama, Thriller',
'Comedy, Drama',
'Comedy, Drama',
'Comedy, Romance',
'Action, Adventure, Drama',
'Crime, Drama, Romance',
'Crime, Drama, Mystery',
'Drama, Horror, Sci-Fi',
'Animation, Comedy',
'Crime, Drama, Mystery',
'Adventure, Drama, Fantasy',
'Crime, Drama',
'Drama, Mystery',
'Crime, Drama, Mystery',
'Drama, Sci-Fi, Thriller',
'Crime, Drama, Mystery',
'Action, Crime, Drama',
'Action, Adventure, Drama',
'Drama, Sci-Fi, Thriller',
'Drama, Fantasy, Horror',
'Drama',
'Comedy',
'Comedy, Drama, Music',
'Animation, Comedy',
```

```
 'Comedy, Crime',
 'Drama, Mystery, Sci-Fi',
 'Action, Adventure, Comedy',
 'Crime, Drama, Mystery',
 'Drama, Fantasy, Mystery',
 'Comedy, Drama, Mystery',
 'Crime, Drama',
 'Drama, Horror, Mystery',
 'Action, Crime, Drama',
 'Comedy, Drama',
 'Drama',
 'Comedy, Drama, Romance',
 'Comedy, Drama, Fantasy',
 'Action, Adventure, Crime',
 'Drama, Sci-Fi, Thriller',
 'Crime, Drama, Mystery',
 'Drama, Mystery, Thriller',
 'Action, Crime, Drama',
 'Drama, Thriller',
 'Action, Adventure, Crime',
 'Crime, Drama, Mystery',
 'Comedy, Drama, Thriller',
 'Crime, Drama, Mystery',
 'Biography, Drama, History',
 'Drama, History, Thriller',
 'Comedy, Crime, Drama',
 'Drama',
 'Adventure, Comedy, Drama',
 'Crime, Drama, Mystery',
 'Comedy, Crime, Drama',
 'Drama, Horror, Mystery']
```

In [28]:

```
try:
    run_tags=driver.find_elements(By.XPATH,'//span[@class="runtime"]')
    for i in run_tags:
        Run_time.append(i.text)
except NoSuchElementException:
    Run_time.append("-")

Run_time
```

Out[28]:

```
['57 min',
 '51 min',
 '44 min',
 '60 min',
 '43 min',
 '59 min',
 '45 min',
 '41 min',
```

'43 min',
'42 min',
'70 min',
'22 min',
'60 min',
'88 min',
'44 min',
'44 min',
'43 min',
'60 min',
'49 min',
'42 min',
'44 min',
'44 min',
'43 min',
'41 min',
'22 min',
'60 min',
'49 min',
'54 min',
'22 min',
'22 min',
'44 min',
'49 min',
'45 min',
'43 min',
'42 min',
'60 min',
'42 min',
'62 min',
'56 min',
'22 min',
'23 min',
'42 min',
'25 min',
'51 min',
'60 min',
'60 min',
'45 min',
'46 min',
'22 min',
'45 min',
'45 min',
'53 min',
'44 min',
'22 min',
'43 min',
'44 min',
'60 min',

```
 '44 min',
 '42 min',
 '44 min',
 '55 min',
 '42 min',
 '42 min',
 '60 min',
 '60 min',
 '41 min',
 '60 min',
 '44 min',
 '22 min',
 '22 min',
 '43 min',
 '60 min',
 '55 min',
 '60 min',
 '45 min',
 '46 min',
 '45 min',
 '53 min',
 '30 min',
 '42 min',
 '45 min',
 '22 min',
 '55 min',
 '45 min',
 '60 min',
 '44 min',
 '55 min',
 '43 min',
 '50 min',
 '60 min',
 '45 min',
 '43 min',
 '58 min',
 '330 min',
 '42 min',
 '42 min',
 '50 min',
 '42 min',
 '45 min',
 '572 min']
```

In [29]:

```
try:
    rating_tags=driver.find_elements(By.XPATH,'//div[@class="ipl-rating-
star small"]')
    for i in rating_tags:
        Ratings.append(i.text)
```

```
except NoSuchElementException:
    Ratings.append("-")
```

```
Ratings
```

Out[29]:

```
['9.2',
 '8.7',
 '8.1',
 '7.5',
 '7.6',
 '8.1',
 '6.5',
 '7.6',
 '7.5',
 '7.5',
 '8.2',
 '8.2',
 '8.8',
 '9.1',
 '8.5',
 '7.4',
 '7.7',
 '8',
 '9.5',
 '8.1',
 '8.4',
 '8.3',
 '8.1',
 '7.7',
 '8.7',
 '7.7',
 '8.8',
 '8.6',
 '8.9',
 '8.3',
 '8.4',
 '8.5',
 '8.3',
 '6.2',
 '7.5',
 '8.2',
 '7.8',
 '8.5',
 '7.9',
 '8.5',
 '9.1',
 '6.5',
 '8',
 '8.7',
```

'8.7',
'7.3',
'8.3',
'8.6',
'7.8',
'7.5',
'7.7',
'8.7',
'6.8',
'8.2',
'8',
'8.3',
'8.8',
'8.7',
'6.7',
'8.3',
'8.3',
'7.3',
'6.8',
'8.4',
'7.4',
'8',
'7.9',
'6.8',
'8.7',
'8.4',
'6.5',
'7.9',
'8.9',
'7.8',
'7.6',
'8.9',
'8.1',
'8.5',
'8.2',
'7.3',
'8.7',
'8.2',
'6.4',
'6.3',
'8.6',
'7.8',
'7.3',
'7.7',
'7.2',
'8.5',
'6.5',
'8.1',
'8.6',

```
 '9.4',
 '7.8',
 '7.4',
 '7.8',
 '8.1',
 '7',
 '8.6']
```

In [31]:

```
try:
    vote_tags=driver.find_elements(By.XPATH,'//span[@name="nv"]')
    for i in vote_tags:
        Votes.append(i.text)
except NoSuchElementException:
    Votes.append("-")

Votes
```

Out[31]:

```
['2,155,232',
 '1,236,966',
 '1,023,895',
 '301,263',
 '260,251',
 '308,901',
 '148,351',
 '320,267',
 '354,886',
 '436,902',
 '493,731',
 '826,571',
 '568,601',
 '947,496',
 '549,393',
 '171,595',
 '330,286',
 '326,398',
 '1,966,120',
 '334,771',
 '457,334',
 '549,907',
 '156,959',
 '155,266',
 '416,525',
 '228,899',
 '439,626',
 '451,611',
 '1,022,586',
 '698,953',
 '423,220',
 '397,281',
```

'140,218',
'126,295',
'180,358',
'157,588',
'234,373',
'514,472',
'218,883',
'447,865',
'548,457',
'66,346',
'201,456',
'513,304',
'406,538',
'83,640',
'297,603',
'252,681',
'232,507',
'220,227',
'276,799',
'736,994',
'134,035',
'349,422',
'260,233',
'565,923',
'577,317',
'477,135',
'62,067',
'113,149',
'348,566',
'76,044',
'107,048',
'246,070',
'100,195',
'100,743',
'53,862',
'151,252',
'386,214',
'330,704',
'108,702',
'257,062',
'591,447',
'108,104',
'132,544',
'568,717',
'111,320',
'246,192',
'94,840',
'23,623',
'149,010',

```
 '170,362',
 '134,487',
 '39,007',
 '303,617',
 '121,919',
 '134,498',
 '76,224',
 '111,274',
 '208,692',
 '30,338',
 '190,306',
 '228,951',
 '792,508',
 '70,707',
 '51,448',
 '63,493',
 '207,027',
 '43,005',
 '256,780']
```

In [32]:

```
try:
    datas_tags=driver.find_elements(By.XPATH,'//h3[@class="lister-item-
header"]')
    for i in datas_tags:
        Name.append(i.text.split('(')[0])
except NoSuchElementException:
    Name.append("-")

Name
```

Out[32]:

```
['1. Game of Thrones ',
 '2. Stranger Things ',
 '3. The Walking Dead ',
 '4. 13 Reasons Why ',
 '5. The 100 ',
 '6. Orange Is the New Black ',
 '7. Riverdale ',
 "8. Grey's Anatomy ",
 '9. The Flash ',
 '10. Arrow ',
 '11. Money Heist ',
 '12. The Big Bang Theory ',
 '13. Black Mirror ',
 '14. Sherlock ',
 '15. Vikings ',
 '16. Pretty Little Liars ',
 '17. The Vampire Diaries ',
 '18. American Horror Story ',
 '19. Breaking Bad ',
```

```
'20. Lucifer ',
'21. Supernatural ',
'22. Prison Break ',
'23. How to Get Away with Murder ',
'24. Teen Wolf ',
'25. The Simpsons ',
'26. Once Upon a Time ',
'27. Narcos ',
'28. Daredevil ',
'29. Friends ',
'30. How I Met Your Mother ',
'31. Suits ',
'32. Mr. Robot ',
'33. The Originals ',
'34. Supergirl ',
'35. Gossip Girl ',
'36. Sense8 ',
'37. Gotham ',
'38. Westworld ',
'39. Jessica Jones ',
'40. Modern Family ',
'41. Rick and Morty ',
'42. Shadowhunters ',
'43. The End of the F***ing World ',
'44. House of Cards ',
'45. Dark ',
'46. Elite ',
'47. Sex Education ',
'48. Shameless ',
'49. New Girl ',
'50. Agents of S.H.I.E.L.D. ',
'51. You ',
'52. Dexter ',
'53. Fear the Walking Dead ',
'54. Family Guy ',
'55. The Blacklist ',
'56. Lost ',
'57. Peaky Blinders ',
'58. House ',
'59. Quantico ',
'60. Orphan Black ',
'61. Homeland ',
'62. Blindspot ',
"63. DC's Legends of Tomorrow ",
"64. The Handmaid's Tale ",
'65. Chilling Adventures of Sabrina ',
'66. The Good Doctor ',
'67. Jane the Virgin ',
'68. Glee ',
```

```
 '69. South Park ',
 '70. Brooklyn Nine-Nine ',
 '71. Under the Dome ',
 '72. The Umbrella Academy ',
 '73. True Detective ',
 '74. The OA ',
 '75. Desperate Housewives ',
 '76. Better Call Saul ',
 '77. Bates Motel ',
 '78. The Punisher ',
 '79. Atypical ',
 '80. Dynasty ',
 '81. This Is Us ',
 '82. The Good Place ',
 '83. Iron Fist ',
 '84. The Rain ',
 '85. Mindhunter ',
 '86. Revenge ',
 '87. Luke Cage ',
 '88. Scandal ',
 '89. The Defenders ',
 '90. Big Little Lies ',
 '91. Insatiable ',
 '92. The Mentalist ',
 '93. The Crown ',
 '94. Chernobyl ',
 '95. iZombie ',
 '96. Reign ',
 '97. A Series of Unfortunate Events ',
 '98. Criminal Minds ',
 '99. Scream: The TV Series ',
 '100. The Haunting of Hill House ']
```

In [33]:

```
print(len(Name),len(Year_span),len(genre),len(Run_time),len(Ratings),len(Votes))
```

```
100 100 100 100 100 100
```

In [34]:

```
import pandas as pd
db=pd.DataFrame({"Name":Name,"Year span":Year_span,"genre":genre,"Run time":Run_time,"Ratings":Ratings,"Votes":Votes})
db
```

Out[34]:

	Name	Year span	genre	Run time	Ratings	Votes
0	1. Game of Thrones	(2011–2019)	Action, Adventure, Drama	57 min	9.2	2,155,232
1	2. Stranger Things	(2016–2024)	Drama, Fantasy, Horror	51 min	8.7	1,236,966

	Name	Year span	genre	Run time	Ratings	Votes
2	3. The Walking Dead	(2010– 2022)	Drama, Horror, Thriller	44 min	8.1	1,023,895
3	4. 13 Reasons Why	(2017– 2020)	Drama, Mystery, Thriller	60 min	7.5	301,263
4	5. The 100	(2014– 2020)	Drama, Mystery, Sci-Fi	43 min	7.6	260,251
...	...	...	...	...	...	...
95	96. Reign	(2013– 2017)	Drama	42 min	7.4	51,448
96	97. A Series of Unfortunate Events	(2017– 2019)	Adventure, Comedy, Drama	50 min	7.8	63,493
97	98. Criminal Minds	(2005–)	Crime, Drama, Mystery	42 min	8.1	207,027
98	99. Scream: The TV Series	(2015– 2019)	Comedy, Crime, Drama	45 min	7	43,005
99	100. The Haunting of Hill House	(2018)	Drama, Horror, Mystery	572 min	8.6	256,780

100 rows × 6 columns

In [ ]:

```
Answer 8
```

In [35]:

```
import selenium
import pandas as pd
import time
from bs4 import BeautifulSoup
from selenium import webdriver
from selenium.common.exceptions import StaleElementReferenceException,
NoSuchElementException
import requests
from selenium.webdriver.common.by import By
```

In [36]:

```
driver=webdriver.Chrome(r"C:\Users\zarna\Downloads\chromedriver_win32
(1)\chromedriver.exe")
```

```
C:\Users\zarna\AppData\Local\Temp\ipykernel_169148\1240552591.py:1:
DeprecationWarning: executable_path has been deprecated, please pass in a
Service object
  driver=webdriver.Chrome(r"C:\Users\zarna\Downloads\chromedriver_win32
(1)\chromedriver.exe")
```

In [42]:

```
driver.get("https://archive.ics.uci.edu/")
```

In [43]:

```
click=driver.find_element(By.XPATH,'/html/body/table[2]/tbody/tr/td/span/b/
a')
click.click()
```

In [44]:

```
Dataset_name=[]
Datatype=[]
```

```
Task=[]
Attribute_type=[]
No_of_instances=[]
No_of_attribute=[]
Year=[]
```

In [45]:

```
data=[]

try:
    data_tags=driver.find_elements(By.XPATH,'//p[@class="normal"]')
    for i in data_tags:
        data.append(i.text)
except NoSuchElementException:
    data.append("-")
```

In [46]:

```
len(data)
```

Out[46]:

```
4364
```

In [47]:

```
data[8]
```

Out[47]:

```
'Abalone'
```

In [49]:

```
data[4361]
```

Out[49]:

```
'2021 '
```

In [52]:

```
datas=[]
datas=data[8:4362]
```

In [53]:

```
len(datas)
```

Out[53]:

```
4354
```

In [55]:

```
Dataset_name=datas[0::7]
Dataset_name
```

Out[55]:

```
['Abalone',
 'Adult',
 'Annealing',
 'Anonymous Microsoft Web Data',
 'Arrhythmia',
 'Artificial Characters',
 'Audiology (Original)',
 'Audiology (Standardized)',
 'Auto MPG',
 'Automobile',
 'Badges',
 'Balance Scale',
 'Balloons',
 'Breast Cancer',
```

```
'Breast Cancer Wisconsin (Original)',
'Breast Cancer Wisconsin (Prognostic)',
'Breast Cancer Wisconsin (Diagnostic)',
'Pittsburgh Bridges',
'Car Evaluation',
'Census Income',
'Chess (King-Rook vs. King-Knight)',
'Chess (King-Rook vs. King-Pawn)',
'Chess (King-Rook vs. King)',
'Chess (Domain Theories)',
'Bach Chorales',
'Connect-4',
'Credit Approval',
'Japanese Credit Screening',
'Computer Hardware',
'Contraceptive Method Choice',
'Covertype',
'Cylinder Bands',
'Dermatology',
'Diabetes',
'DGP2 - The Second Data Generation Program',
'Document Understanding',
'EBL Domain Theories',
'Echocardiogram',
'Ecoli',
'Flags',
'Function Finding',
'Glass Identification',
"Haberman's Survival",
'Hayes-Roth',
'Heart Disease',
'Hepatitis',
'Horse Colic',
'ICU',
'Image Segmentation',
'Internet Advertisements',
'Ionosphere',
'Iris',
'ISOLET',
'Kinship',
'Labor Relations',
'LED Display Domain',
'Lenses',
'Letter Recognition',
'Liver Disorders',
'Logic Theorist',
'Lung Cancer',
'Lymphography',
'Mechanical Analysis',
```

'Meta-data',
'Mobile Robots',
'Molecular Biology (Promoter Gene Sequences)',
'Molecular Biology (Protein Secondary Structure)',
'Molecular Biology (Splice-junction Gene Sequences)',
"MONK's Problems",
'Moral Reasoner',
'Multiple Features',
'Mushroom',
'Musk (Version 1)',
'Musk (Version 2)',
'Nursery',
'Othello Domain Theory',
'Page Blocks Classification',
'Optical Recognition of Handwritten Digits',
'Pen-Based Recognition of Handwritten Digits',
'Post-Operative Patient',
'Primary Tumor',
'Prodigy',
'Qualitative Structure Activity Relationships',
'Quadruped Mammals',
'Servo',
'Shuttle Landing Control',
'Solar Flare',
'Soybean (Large)',
'Soybean (Small)',
'Challenger USA Space Shuttle O-Ring',
'Low Resolution Spectrometer',
'Spambase',
'SPECT Heart',
'SPECTF Heart',
'Sponge',
'Statlog Project',
'Student Loan Relational',
'Teaching Assistant Evaluation',
'Tic-Tac-Toe Endgame',
'Thyroid Disease',
'Trains',
'University',
'Congressional Voting Records',
'Water Treatment Plant',
'Waveform Database Generator (Version 1)',
'Waveform Database Generator (Version 2)',
'Wine',
'Yeast',
'Zoo',
'Undocumented',
'Twenty Newsgroups',
'Australian Sign Language signs',

'Australian Sign Language signs (High Quality)',
'US Census Data (1990)',
'Census-Income (KDD)',
'Coil 1999 Competition Data',
'Corel Image Features',
'E. Coli Genes',
'EEG Database',
'El Nino',
'Entree Chicago Recommendation Data',
'CMU Face Images',
'Insurance Company Benchmark (COIL 2000)',
'Internet Usage Data',
'IPUMS Census Database',
'Japanese Vowels',
'KDD Cup 1998 Data',
'KDD Cup 1999 Data',
'M. Tuberculosis Genes',
'Movie',
'MSNBC.com Anonymous Web Data',
'NSF Research Award Abstracts 1990-2003',
'Pioneer-1 Mobile Robot Data',
'Pseudo Periodic Synthetic Time Series',
'Reuters-21578 Text Categorization Collection',
'Robot Execution Failures',
'Synthetic Control Chart Time Series',
'Syskill and Webert Web Page Ratings',
'UNIX User Data',
'Volcanoes on Venus - JARtool experiment',
'Statlog (Australian Credit Approval)',
'Statlog (German Credit Data)',
'Statlog (Heart)',
'Statlog (Landsat Satellite)',
'Statlog (Image Segmentation)',
'Statlog (Shuttle)',
'Statlog (Vehicle Silhouettes)',
'Connectionist Bench (Nettalk Corpus)',
'Connectionist Bench (Sonar, Mines vs. Rocks)',
'Connectionist Bench (Vowel Recognition - Deterding Data)',
'Economic Sanctions',
'Protein Data',
'Cloud',
'CalIt2 Building People Counts',
'Dodgers Loop Sensor',
'Poker Hand',
'MAGIC Gamma Telescope',
'UJI Pen Characters',
'Mammographic Mass',
'Forest Fires',
'Reuters Transcribed Subset',

'Bag of Words',
'Concrete Compressive Strength',
'Hill-Valley',
'Arcene',
'Dexter',
'Dorothea',
'Gisette',
'Madelon',
'Ozone Level Detection',
'Abscisic Acid Signaling Network',
'Parkinsons',
'Character Trajectories',
'Blood Transfusion Service Center',
'UJI Pen Characters (Version 2)',
'Semeion Handwritten Digit',
'SECOM',
'Plants',
'Libras Movement',
'Concrete Slump Test',
'Communities and Crime',
'Acute Inflammations',
'Wine Quality',
'URL Reputation',
'p53 Mutants',
'Parkinsons Telemonitoring',
'Demospongiae',
'Opinosis Opinion ⁄ Review',
'Breast Tissue',
'Cardiotocography',
'Wall-Following Robot Navigation Data',
'Spoken Arabic Digit',
'Localization Data for Person Activity',
'AutoUniv',
'Steel Plates Faults',
'MiniBooNE particle identification',
'YearPredictionMSD',
'PEMS-SF',
'OpinRank Review Dataset',
'Relative location of CT slices on axial axis',
'Online Handwritten Assamese Characters Dataset',
'PubChem Bioassay Data',
'Record Linkage Comparison Patterns',
'Communities and Crime Unnormalized',
'Vertebral Column',
'EMG Physical Action Data Set',
'Vicon Physical Action Data Set',
'Amazon Commerce reviews set',
'Amazon Access Samples',
'Reuter_50_50',

```
 'Farm Ads',
 'DBWorld e-mails',
 'KEGG Metabolic Relation Network (Directed)',
 'KEGG Metabolic Reaction Network (Undirected)',
 'Bank Marketing',
 'YouTube Comedy Slam Preference Data',
 'Gas Sensor Array Drift Dataset',
 'ILPD (Indian Liver Patient Dataset)',
 'OPPORTUNITY Activity Recognition',
 'Nomao',
 'SMS Spam Collection',
 'Skin Segmentation',
 'Planning Relax',
 'PAMAP2 Physical Activity Monitoring',
 'Restaurant & consumer data',
 'CNAE-9',
 'Individual household electric power consumption',
 'seeds',
 'Northix',
 'QtyT40I10D100K',
 'Legal Case Reports',
 'Human Activity Recognition Using Smartphones',
 'One-hundred plant species leaves data set',
 'Energy efficiency',
 'Yacht Hydrodynamics',
 'Fertility',
 'Daphnet Freezing of Gait',
 '3D Road Network (North Jutland, Denmark)',
 'ISTANBUL STOCK EXCHANGE',
 'Buzz in social media',
 'First-order theorem proving',
 'Wearable Computing: Classification of Body Postures and Movements (PUC-
Rio)',
 'Gas sensor arrays in open sampling settings',
 'Climate Model Simulation Crashes',
 'MicroMass',
 'QSAR biodegradation',
 'BLOGGER',
 'Daily and Sports Activities',
 'User Knowledge Modeling',
 'Reuters RCV1 RCV2 Multilingual, Multiview Text Categorization Test
collection',
 'NYSK',
 'Turkiye Student Evaluation',
 "ser Knowledge Modeling Data (Students' Knowledge Levels on DC Electrical
Machines)",
 'EEG Eye State',
 'Physicochemical Properties of Protein Tertiary Structure',
 'seismic-bumps',
```

```
 'banknote authentication',
 'USPTO Algorithm Challenge, run by NASA-Harvard Tournament Lab and
TopCoder Problem: Pat',
 'YouTube Multiview Video Games Dataset',
 'Gas Sensor Array Drift Dataset at Different Concentrations',
 'Activities of Daily Living (ADLs) Recognition Using Binary Sensors',
 'SkillCraft1 Master Table Dataset',
 'Weight Lifting Exercises monitored with Inertial Measurement Units',
 'SML2010',
 'Bike Sharing Dataset',
 'Predict keywords activities in a online social media',
 'Thoracic Surgery Data',
 'EMG dataset in Lower Limb',
 'SUSY',
 'HIGGS',
 'Qualitative_Bankruptcy',
 'LSVT Voice Rehabilitation',
 'Dataset for ADL Recognition with Wrist-worn Accelerometer',
 'Wilt',
 'User Identification From Walking Activity',
 'Activity Recognition from Single Chest-Mounted Accelerometer',
 'Leaf',
 'Dresses_Attribute_Sales',
 'Tamilnadu Electricity Board Hourly Readings',
 'Airfoil Self-Noise',
 'Wholesale customers',
 'Twitter Data set for Arabic Sentiment Analysis',
 'Combined Cycle Power Plant',
 'Urban Land Cover',
 'Diabetes 130-US hospitals for years 1999-2008',
 'Bach Choral Harmony',
 'StoneFlakes',
 'Tennis Major Tournament Match Statistics',
 'Parkinson Speech Dataset with Multiple Types of Sound Recordings',
 'Gesture Phase Segmentation',
 'Perfume Data',
 'BlogFeedback',
 'REALDISP Activity Recognition Dataset',
 'Newspaper and magazine images segmentation dataset',
 'AAAI 2014 Accepted Papers',
 'Gas sensor array under flow modulation',
 'Gas sensor array exposed to turbulent gas mixtures',
 'UJIIndoorLoc',
 'Sentence Classification',
 'Dow Jones Index',
 'sEMG for Basic Hand movements',
 'AAAI 2013 Accepted Papers',
 'Geographical Original of Music',
 'Condition Based Maintenance of Naval Propulsion Plants',
```

```
 'Grammatical Facial Expressions',
 'NoisyOffice',
 'MHEALTH Dataset',
 'Student Performance',
 'ElectricityLoadDiagrams20112014',
 'Gas sensor array under dynamic gas mixtures',
 'microblogPCU',
 'Firm-Teacher_Clave-Direction_Classification',
 'Dataset for Sensorless Drive Diagnosis',
 'TV News Channel Commercial Detection Dataset',
 'Phishing Websites',
 'Greenhouse Gas Observing Network',
 'Diabetic Retinopathy Debrecen Data Set',
 'HIV-1 protease cleavage',
 'Sentiment Labelled Sentences',
 'Online News Popularity',
 'Forest type mapping',
 'wiki4HE',
 'Online Video Characteristics and Transcoding Time Dataset',
 'Chronic_Kidney_Disease',
 'Machine Learning based ZZAlpha Ltd. Stock Recommendations 2012-2014',
 'Folio',
 'Taxi Service Trajectory - Prediction Challenge, ECML PKDD 2015',
 'Cuff-Less Blood Pressure Estimation',
 'Smartphone-Based Recognition of Human Activities and Postural
Transitions',
 'Mice Protein Expression',
 'UJIIndoorLoc-Mag',
 'Heterogeneity Activity Recognition',
 'Educational Process Mining (EPM): A Learning Analytics Data Set',
 'HEPMASS',
 'Indoor User Movement Prediction from RSS data',
 'Open University Learning Analytics dataset',
 'default of credit card clients',
 'Mesothelioma’s disease data set',
 'Online Retail',
 'SIFT10M',
 'GPS Trajectories',
 'Detect Malacious Executable(AntiVirus)',
 'Occupancy Detection',
 'Improved Spiral Test Using Digitized Graphics Tablet for Monitoring
Parkinson’s Disease',
 'News Aggregator',
 'Air Quality',
 'Twin gas sensor arrays',
 'Gas sensors for home activity monitoring',
 'Facebook Comment Volume Dataset',
 'Smartphone Dataset for Human Activity Recognition (HAR) in Ambient
Assisted Living (AAL)',
```

```
 'Polish companies bankruptcy data',
 'Activity Recognition system based on Multisensor data fusion (AReM)',
 'Dota2 Games Results',
 'Facebook metrics',
 'UbiqLog (smartphone lifelogging)',
 'NIPS Conference Papers 1987-2015',
 'HTRU2',
 'Drug consumption (quantified)',
 'Appliances energy prediction',
 'Miskolc IIS Hybrid IPS',
 'KDC-4007 dataset Collection',
 'Geo-Magnetic field and WLAN dataset for indoor localisation from
wristband and smartphone',
 'DrivFace',
 'Website Phishing',
 'YouTube Spam Collection',
 'Beijing PM2.5 Data',
 'Cargo 2000 Freight Tracking and Tracing',
 'Cervical cancer (Risk Factors)',
 'Quality Assessment of Digital Colposcopies',
 'KASANDR',
 'FMA: A Dataset For Music Analysis',
 'Air quality',
 'Epileptic Seizure Recognition',
 'Devanagari Handwritten Character Dataset',
 'Stock portfolio performance',
 'MoCap Hand Postures',
 'Early biomarkers of Parkinson's disease based on natural connected
speech',
 'Data for Software Engineering Teamwork Assessment in Education Setting',
 'PM2.5 Data of Five Chinese Cities',
 'Parkinson Disease Spiral Drawings Using Digitized Graphics Tablet',
 'Sales_Transactions_Dataset_Weekly',
 'Las Vegas Strip',
 'Eco-hotel',
 'MEU-Mobile KSD',
 'Crowdsourced Mapping',
 'gene expression cancer RNA-Seq',
 'Hybrid Indoor Positioning Dataset from WiFi RSSI, Bluetooth and
magnetometer',
 'chestnut - LARVIC',
 'Burst Header Packet (BHP) flooding attack on Optical Burst Switching
(OBS) Network',
 'Motion Capture Hand Postures',
 'Anuran Calls (MFCCs)',
 'TTC-3600: Benchmark dataset for Turkish text categorization',
 'Gastrointestinal Lesions in Regular Colonoscopy',
 'Daily Demand Forecasting Orders',
 'Paper Reviews',
```

```
 'extention of Z-Alizadeh sani dataset',
 'Z-Alizadeh Sani',
 'Dynamic Features of VirusShare Executables',
 'IDA2016Challenge',
 'DSRC Vehicle Communications',
 'Mturk User-Perceived Clusters over Images',
 'Character Font Images',
 'DeliciousMIL: A Data Set for Multi-Label Multi-Instance Learning with
Instance Labels',
 'Autistic Spectrum Disorder Screening Data for Children',
 'Autistic Spectrum Disorder Screening Data for Adolescent',
 'APS Failure at Scania Trucks',
 'Wireless Indoor Localization',
 'HCC Survival',
 'CSM (Conventional and Social Media Movies) Dataset 2014 and 2015',
 'University of Tehran Question Dataset 2016 (UTQD.2016)',
 'Autism Screening Adult',
 'Activity recognition with healthy older people using a batteryless
wearable sensor',
 'Immunotherapy Dataset',
 'Cryotherapy Dataset',
 'OCT data & Color Fundus Images of Left & Right Eyes',
 'Discrete Tone Image Dataset',
 'News Popularity in Multiple Social Media Platforms',
 'Ultrasonic flowmeter diagnostics',
 'ICMLA 2014 Accepted Papers Data Set',
 'BLE RSSI Dataset for Indoor localization and Navigation',
 'Container Crane Controller Data Set',
 'Residential Building Data Set',
 'Health News in Twitter',
 'chipseq',
 'SGEMM GPU kernel performance',
 'Repeat Consumption Matrices',
 'detection_of_IoT_botnet_attacks_N_BaIoT',
 'Absenteeism at work',
 'SCADI',
 'Condition monitoring of hydraulic systems',
 'Carbon Nanotubes',
 'Optical Interconnection Network',
 'Sports articles for objectivity analysis',
 'Breast Cancer Coimbra',
 'GNFUV Unmanned Surface Vehicles Sensor Data',
 'Dishonest Internet users Dataset',
 'Victorian Era Authorship Attribution',
 'Simulated Falls and Daily Living Activities Data Set',
 'Multimodal Damage Identification for Humanitarian Computing',
 'EEG Steady-State Visual Evoked Potential Signals',
 'Roman Urdu Data Set',
 'Avila',
```

```
 'PANDOR',
 'Drug Review Dataset (Druglib.com)',
 'Drug Review Dataset (Drugs.com)',
 'Physical Unclonable Functions',
 'Superconductivty Data',
 'WESAD (Wearable Stress and Affect Detection)',
 'GNFUV Unmanned Surface Vehicles Sensor Data Set 2',
 'Student Academics Performance',
 'Online Shoppers Purchasing Intention Dataset',
 'PMU-UD',
 "Parkinson's Disease Classification",
 'Electrical Grid Stability Simulated Data',
 'Caesarian Section Classification Dataset',
 'BAUM-1',
 'BAUM-2',
 'Audit Data',
 'BuddyMove Data Set',
 'Real estate valuation data set',
 'Early biomarkers of Parkinson’s disease based on natural connected speech
Data Set',
 'Somerville Happiness Survey',
 '2.4 GHZ Indoor Channel Measurements',
 'EMG data for gestures',
 'Parking Birmingham',
 'Behavior of the urban traffic of the city of Sao Paulo in Brazil',
 'Travel Reviews',
 'Tarvel Review Ratings',
 'Rice Leaf Diseases',
 'Gas sensor array temperature modulation',
 'Facebook Live Sellers in Thailand',
 'Parkinson Dataset with replicated acoustic features',
 'Metro Interstate Traffic Volume',
 'Query Analytics Workloads Dataset',
 'Wave Energy Converters',
 'PPG-DaLiA',
 'Alcohol QCM Sensor Dataset',
 'Divorce Predictors data set',
 'Incident management process enriched event log',
 'Opinion Corpus for Lebanese Arabic Reviews (OCLAR)',
 'MEx',
 'Beijing Multi-Site Air-Quality Data',
 'Online Retail II',
 'Hepatitis C Virus (HCV) for Egyptian patients',
 'QSAR fish toxicity',
 'QSAR aquatic toxicity',
 'Human Activity Recognition from Continuous Ambient Sensor Data',
 'WISDM Smartphone and Smartwatch Activity and Biometrics Dataset',
 'QSAR oral toxicity',
 'QSAR androgen receptor',
```

```
 'QSAR Bioconcentration classes dataset',
 'QSAR fish bioconcentration factor (BCF)',
 'A study of Asian Religious and Biblical Texts',
 'Real-time Election Results: Portugal 2019',
 'Bias correction of numerical prediction model temperature forecast',
 'Bar Crawl: Detecting Heavy Drinking',
 'Kitsune Network Attack Dataset',
 'Shoulder Implant X-Ray Manufacturer Classification',
 'Speaker Accent Recognition',
 'Heart failure clinical records',
 'Deepfakes: Medical Image Tamper Detection',
 'selfBACK',
 'South German Credit',
 'Exasens',
 'Swarm Behaviour',
 'Crop mapping using fused optical-radar data set',
 'BitcoinHeistRansomwareAddressDataset',
 'Facebook Large Page-Page Network',
 'Amphibians',
 'Early stage diabetes risk prediction dataset.',
 'Turkish Spam V01',
 'Stock keeping units',
 'Demand Forecasting for a store',
 'Detect Malware Types',
 'Wave Energy Converters',
 'Youtube cookery channels viewers comments in Hinglish',
 'Pedestrian in Traffic Dataset',
 'Cervical Cancer Behavior Risk',
 'Sattriya_Dance_Single_Hand_Gestures Dataset',
 'Divorce Predictors data set',
 '3W dataset',
 'Malware static and dynamic features VxHeaven and Virus Total',
 'Internet Firewall Data',
 'User Profiling and Abusive Language Detection Dataset',
 'Estimation of obesity levels based on eating habits and physical
condition',
 'Rice (Cammeo and Osmancik)',
 'Vehicle routing and scheduling problems',
 'Algerian Forest Fires Dataset',
 'Breath Metabolomics',
 'Horton General Hospital',
 'UrbanGB, urban road accidents coordinates labelled by the urban center',
 'Gas Turbine CO and NOx Emission Data Set',
 'Activity recognition using wearable physiological measurements',
 'clickstream data for online shopping',
 'CNNpred: CNN-based stock market prediction using a diverse set of
variables',
 'Apartment for rent classified',
 ': Simulated Data set of Iraqi tourism places',
```

'Nasarian CAD Dataset',
'Monolithic Columns in Troad and Mysia Region',
'Bar Crawl: Detecting Heavy Drinking',
'Seoul Bike Sharing Demand',
'Person Classification Gait Data',
'Shill Bidding Dataset',
'Iranian Churn Dataset',
'Unmanned Aerial Vehicle (UAV) Intrusion Detection',
'Bone marrow transplant: children',
'Exasens',
'COVID-19 Surveillance',
'Refractive errors',
'Shoulder Implant X-Ray Manufacturer Classification',
'CLINC150',
'HCV data',
'Taiwanese Bankruptcy Prediction',
'South German Credit (UPDATE)',
'IIWA14-R820-Gazebo-Dataset-10Trajectories',
'Guitar Chords finger positions',
'Russian Corpus of Biographical Texts',
'Codon usage',
'Intelligent Media Accelerometer and Gyroscope (IM-AccGyro) Dataset',
'Myocardial infarction complications',
'Hungarian Chickenpox Cases',
'Simulated data for survival modelling',
'Student Performance on an entrance examination',
'Chemical Composition of Ceramic Samples',
'Labeled Text Forum Threads Dataset',
'Stock keeping units',
'BLE RSSI dataset for Indoor localization',
'Basketball dataset',
'GitHub MUSAE',
'Anticancer peptides',
'Monolithic Columns in Troad and Mysia Region',
'Gender by Name',
'Iranian Churn Dataset',
'Unmanned Aerial Vehicle (UAV) Intrusion Detection',
'Shoulder Implant Manufacture Classification',
'LastFM Asia Social Network',
'Wheat kernels',
'Productivity Prediction of Garment Employees',
'Multi-view Brain Networks',
'LastFM Asia Social Network',
'Wisesight Sentiment Corpus',
'AI4I 2020 Predictive Maintenance Dataset',
'Dry Bean Dataset',
'in-vehicle coupon recommendation',
'Gait Classification',
'Wikipedia Math Essentials',

```
 'Wikipedia Math Essentials',
 'Synchronous Machine Data Set',
 'Average Localization Error (ALE) in sensor node localization process in
WSNs',
 '9mers from cullpdb',
 'TamilSentiMix',
 'Accelerometer',
 'Synchronous Machine Data Set',
 'Pedal Me Bicycle Deliveries',
 'Turkish Headlines Dataset',
 'Secondary Mushroom Dataset',
 'Power consumption of Tetouan city',
 'Raisin Dataset',
 'Steel Industry Energy Consumption Dataset',
 'Gender Gap in Spanish WP',
 'Non verbal tourists data',
 'Roman Urdu Sentiment Analysis Dataset (RUSAD)',
 'TUANDROMD ( Tezpur University Android Malware Dataset)',
 'Higher Education Students Performance Evaluation Dataset',
 'Risk Factor prediction of Chronic Kidney Disease',
 'Lab Test',
 'Shoulder Implant Manufacture Classification',
 'Rocket League Skillshots Data Set',
 'Sepsis survival minimal clinical records',
 'Water Quality Prediction',
 'Traffic Flow Forecasting',
 'sentiment analysis in Saudi Arabia about distance education during Covid-
19',
 'Kain Tradisional Sambas',
 'Image Recognition Task Execution Times in Mobile Edge Computing',
 'REWEMA',
 'REJAFADA',
 'Steel Industry Energy Consumption Dataset',
 'Influenza outbreak event prediction via Twitter data',
 'Turkish Music Emotion Dataset',
 'Maternal Health Risk Data Set',
 'Room Occupancy Estimation',
 'Image Recognition Task Execution Times in Mobile Edge Computing']
```

In [56]:

```
Datatype=datas[1::7]
Task=datas[2::7]
Attribute_type=datas[3::7]
No_of_instances=datas[4::7]
No_of_attribute=datas[5::7]
Year=datas[6::7]
```

In [57]:

```
print(len(Datatype),len(Task),len(Attribute_type),len(No_of_instances),len(
No_of_attribute),len(Year),len(Dataset_name))
622 622 622 622 622 622 622
```

In [59]:

```
import pandas as pd
db=pd.DataFrame({"Dataset name":Dataset_name,"Data
Type":Datatype,'Task':Task,"Attribute type":Attribute_type,"No of
instances":No_of_instances,"No_of_attribute":No_of_attribute,"Year":Year})
db
```

Out[59]:

	Dataset name	Data Type	Task	Attribute type	No of instances	No_of_attribute	Year
0	Abalone	Multivariate	Classification	Categorical, Integer, Real	4177	8	1995
1	Adult	Multivariate	Classification	Categorical, Integer	48842	14	1996
2	Annealing	Multivariate	Classification	Categorical, Integer, Real	798	38	
3	Anonymous Microsoft Web Data		Recommender-Systems	Categorical	37711	294	1998
4	Arrhythmia	Multivariate	Classification	Categorical, Integer, Real	452	279	1998
...	...	...	...	...	...	...	...
617	Influenza outbreak event prediction via Twitte...	Multivariate	Classification	Integer, Real	75840	525	2020
618	Turkish Music Emotion Dataset	Multivariate	Classification	Integer, Real	400	50	2020
619	Maternal Health Risk Data Set		Classification		1014	7	2020
620	Room Occupancy Estimation	Multivariate, Time-Series	Classification	Real	10129	16	2021
621	Image Recognition Task Execution Times in Mobi...	Univariate	Regression	Real	4000	2	2021

622 rows × 7 columns

In [ ]:

```
Answer 9
```

In [60]:

```
import selenium
```

```
import pandas as pd
import time
from bs4 import BeautifulSoup
from selenium import webdriver
from selenium.common.exceptions import StaleElementReferenceException,
NoSuchElementException
import requests
from selenium.webdriver.common.by import By
```

In [61]:

```
driver=webdriver.Chrome(r"C:\Users\zarna\Downloads\chromedriver_win32
(1)\chromedriver.exe")
```

```
C:\Users\zarna\AppData\Local\Temp\ipykernel_169148\1240552591.py:1:
DeprecationWarning: executable_path has been deprecated, please pass in a
Service object
  driver=webdriver.Chrome(r"C:\Users\zarna\Downloads\chromedriver_win32
(1)\chromedriver.exe")
```

In [65]:

```
driver.get("https:www.naukri.com/hr-recruiters-consultants")
```

In [69]:

```
skill=driver.find_element(By.CLASS_NAME,'sugInp')
skill.send_keys('Data Science')
```

In [70]:

```
click=driver.find_element(By.XPATH,'/html/body/div[2]/div[2]/div[1]/div[1]/
form/div[1]/button')
click.click()
```

In [71]:

```
Name=[]
Designation=[]
Company=[]
Skills=[]
Location=[]
```

In [72]:

```
try:
    name_tags=driver.find_elements(By.XPATH,'//span[@class="fl ellipsis"]')
    for i in name_tags:
        Name.append(i.text)
except NoSuchElementException:
    Name.append('-')

try:
    Designation_tags=driver.find_elements(By.XPATH,'//span[@class="ellipsis
clr"]')
    for i in Designation_tags:
        Designation.append(i.text)
except NoSuchElementException:
    Designation.append('-')
```

In [78]:

```
Name
```

Out[78]:

```
['Aakash Harit',
```

'shravan Kumar Gaddam',
'MARSIAN Technologies LLP',
'Anik Agrawal',
'subhas patel',
'Abhishek - Only Analytics Hiring - India and',
'Institute for Financial Management and Resear',
'Balu Ramesh',
'Asif Lucknowi',
'InstaFinancials',
'Kalpana Dumpala',
'Mubarak',
'Kushal Rastogi',
'Priyanka Akiri',
'Kapil Devang',
'Mahesh Babu Channa',
'Vaishnavi Kudalkar',
'Sakshi Chhikara',
'Ruchi Dhote',
'Manisha Yadav',
'Riya Rajesh',
'Rashmi Bhattacharjee',
'Faizan Kareem',
'Rithika dadwal',
'Sandhya Khandagale',
'Shaun Rao',
'Deeparchi Sharma',
'Azahar Shaikh',
'Manas',
'kumar',
'Sunil Vedula',
'Rajat Kumar',
'Dhruv Dev Dubey',
'Avnish Mishra',
'Jayanth N',
'Avodha',
'Priya Khare',
'Amit Sharma',
'Kanan',
'Shashikant Chaudhary',
'Brad',
'Rutuja Pawar',
'Madhusudhan Sridhar',
'Ankit Sinha',
'Gaurav Chouhan',
'Rashi Kacker',
'Ashwini',
'Balaji Kolli',
'Rajani Nagaraj',
'ROHIT Kumar']

In [79]:

```
Designation
```

Out[79]:

```
['HR Manager',
 'Company Recruiter',
 'Company HR',
 'Company Recruiter',
 'Founder CEO',
 'Recruitment Lead Consultant',
 'Programme Manager',
 'HR Administrator',
 'Director',
 'Human Resource',
 'Executive Hiring',
 'Company HR',
 'Company HR',
 'HR Manager',
 'HR Manager',
 'HR Team Lead',
 'HR Executive',
 'Assistant Manager HR',
 'Senior Executive Talent Acquisition',
 'HR Executive',
 'Manager Talent Acquisition',
 'HR Head',
 'HR MANAGER',
 'HR Recruiter',
 'HR Recruiter',
 'Manager Human Resources',
 'Company Recruiter',
 'Company Recruiter',
 'Lead Talent acquisition',
 'Proprietor',
 'CEO',
 'Founder CEO',
 'Company Recruitment Head',
 'Senior Technical &amp;amp; Corporate Recruiter',
 'Project Manager',
 'Business Development Associate',
 'Senior Manager',
 'Consultant',
 'senior technology instructor',
 'HR Recruiter/HR Excutive',
 'Manager, Technical Recruiting',
 'Technical Recruiter',
 'Erp Implementer',
 'Head Analytics',
 'Chief Technical Officer',
 'Sr Product Manager',
```

```
 'Director Global Delivery',
 'Co Founder',
 'HR Manager',
 'Architect']
```

In [80]:

```
try:

Location_tags=driver.find_elements(By.XPATH,'//small[@class="ellipsis"]')
    for i in Location_tags:
        Location.append(i.text)
except NoSuchElementException:
    Location.append('-')

Location
```

Out[80]:

```
['Delhi',
 'Hyderabad / Secunderabad',
 'Pune',
 'Ahmedabad',
 'UK - (london)',
 'Vadodara / Baroda',
 'Chennai',
 'Trivandrum',
 'Indore',
 'Bengaluru / Bangalore',
 'Hyderabad / Secunderabad',
 'Bengaluru / Bangalore',
 'Mumbai',
 'Hyderabad',
 'Bhopal',
 'Hyderabad / Secunderabad',
 'Mumbai',
 'Chandigarh',
 'Pune',
 'Navi Mumbai',
 'Cochin',
 'Delhi',
 'Hyderabad / Secunderabad',
 'Pune',
 'Pune',
 'Pune',
 'Gurgaon',
 'Pune',
 'Bengaluru / Bangalore',
 'Bengaluru / Bangalore',
 'Delhi',
 'Bengaluru / Bangalore',
 'Noida',
 'Mysoru / Mysore',
```

```
 'Hyderabad / Secunderabad',
 'Bengaluru / Bangalore',
 'New Delhi',
 'Chennai',
 'Aligarh',
 'Salt Lake City',
 'Pune',
 'Bengaluru / Bangalore',
 'Mumbai',
 'Indore',
 'Bengaluru / Bangalore',
 'MYSORE',
 'Hyderabad / Secunderabad',
 'Bengaluru / Bangalore',
 'Mumbai']
```

In [81]:

```
try:
    Skills_tags=driver.find_elements(By.XPATH,'//div[@class="hireSec
highlightable"]')
    for i in Skills_tags:
        Skills.append(i.text)
except NoSuchElementException:
    Skills.append('-')

Skills
```

Out[81]:

```
['Classic ASP Developer, Internet Marketing Professional, Data Science SME,
Content Writers, SEO Professional, Revenue Professional',
 '.Net, Java, Data Science, Linux Administration, Sql Server Development,
Winforms, Wcf Services, Wpf, Telecom Engineering, Technical Management,
Software',
 'Data Science, Artificial Intelligence, Machine Learning, Business
Analytics, Deep Learning, statistics, Data Analytics, Data Analysis,
support vector machine',
 'Mean Stack, javascript, angularjs, mongodb, Web Services, rest, express,
Node.js, Big Data, iot, Data Science, Cloud Computing, saas, Aws',
 'Hadoop, Spark, Digital Strategy, Data Architecture, Command Center, Cdp,
Dmp, Kafka, Data Science, Data Analysis, Big Data Analytics, Real Time
Analysis, SQL',
 'Analytics, Business Intelligence, Business Analytics, Predictive
Modeling, Predictive Analytics, Data Science, Data Analysis, Data
Analytics, Big Data, Big',
 'Data Science',
 'Machine Learning, algorithms, Go Getter, Computer Science, spark, Big
Data, hdfs, sql, cassandra, hadoop, python, scala, java, Data Science,
Front End',
 'Technical Training, Software Development, Presentation Skills, B.tech,
M.tech, B.e., mca, msc, Computer Science, freshers, jobs in indore, Data
Science, itil',
```

'Software Development, It Sales, Account Management, Data Analysis, Customer Service, Sr, Software Engineering, Mvc, Ajax, Asp.net, Html, C#, Javascript',
'Qa, Ui/ux, Java Developer, Java Architect, C++/qt, Php, Lamp, Api, J2ee, Java, Soa, Esb, Middleware, Bigdata Achitect, Hadoop Architect, Deep',
'Business Intelligence, Data Warehousing, Data Science, Business Analytics, Customer Support, Business Reporting, Bi',
'Office Administration, Hr Administration, telecalling, client relationship management, Client Acquisition, Sales, Reception, HR, Recruitment, Onboarding, Human',
'Oracle Dba, Data Science, Data Warehousing, ETL, Jupyter, Numpy, Data Transformation, Snowflake, Teradata, Python, Data Manipulation, Relational Databases',
'Big Data, Hadoop, Data Analytics, Data Science',
'Social Media, digital media maketing, seo, smm, smo, sem, Content Wirting, social media marketing, social media manager, digital media marketing manager',
'Data Science, Python, Data Analytics',
'React.js, Data Science, Java, Front End, Business Analytics, Backend, Tableau, Python, Qa Testing, Automation Testing',
'Qlikview, Qlik Sense, Microsoft Azure, Power Bi, Data Science, Machine Learning',
'Telecalling, Client Interaction, Marketing, Research, Web Development, Social Media Marketing, Data Entry Operation, Excel, Ms Office, Invoicing',
'Data Science',
'Corporate Sales, Software Development, Software Sales, Marketing, Creative Designing, Corporate Planning, Senior Management, Crm, Client Relationship',
'Data Analytics, Data Science, Machine Learning, Deep Learning, Nlp, Data Mining, Python, R, Database Administration, Text Mining',
'Data Science, Machine Learning, Python, R, Deep Learning, Big Data, Hadoop',
'Big Data, Data Science, Artificial Intelligence, Hadoop, Ui Development, Php, Freelancing, .Net, Software Testing, Sap, Leadership Hiring',
'Java, Net, Angularjs, Hr, Infrastructure, Management, Project Management, Business Analysis, Data Science, Information Technology, Technology',
'Research, Digital Marketing, Analytics, Software Development, Data Science, Consulting',
'Data Science, Artificial Intelligence, Machine Learning, Data Analytics',
'Software Architecture, Vp Engineering, Product Management, analytics, Data Science, Node.js, Principal Engineer, Big Data, python, angularjs, React.js',
'Data Science, Hadoop, Rpas, Devops, Python, Aws, Teaching, Big Data',
'Signal Processing, Machine Learning, Neural Networks, Data Science, Predictive Analytics, Time Series Analysis, Data Visualization, Technical Leadership, Data',
'Web Technologies, Project Management, Software Architecture, Data Science, Object Oriented Programming, Computer Science, Electrical Engineering, Architecture',

```
 'Server Administartion, Verilog, Vhdl, Digital Marketing, Market Research,
Property Research, Legal, It And Non It Recruitment, Logistics, Supply
Chain, Bfsi',
 'Data Science, Machine Learning, Deep Learning, Accounting, Statistics',
 'Data Analytics, Managed Services, Team Leading, python, Machine Learning,
Google Analytics, Dmp, Aws, Campaign Analytics, Digital Campaigns,
Audience',
 'Ethical Hacking, Security Operations Center, SOC, Managed Services, Data
Science, Machine Learning, Artificial Intelligence, Operations Research,
Education',
 'Data Science, Artificial Intelligence, analytics, Business Intelligence,
python, tableau, Power Bi, qlikview, sql, Data Warehousing, Data
Visualization',
 'Machine Learning, Artificial Intelligence, Data Science, Software
Engineering, Software Development, Graduate Engineer Trainee, Fresher, Data
Analytics, Java',
 'C, C++, Artificial Intelligence, Python, Php, Web Development, Matlab,
Data Science, Augmented Reality, C C++',
 'Relationship Management, Retail Sales, Private Banking, Mutual Funds,
NISM, Equity, Finance, Financial Products, Financial Services, Verbal,
Written',
 'Data Science, Software Engineering',
 'Data Science, Big Data Analytics, Digital Marketing, Content Writing, Ui
Development, Database Development, Qa Automation, Python, Project
Management',
 'Data Science, Recruitment, Salary',
 'B.Tech, Tableau, Statistics, R, Analytics, Time Series, Data Science,
Business Solutions, SQL, Technical Skills, SSAS, SQL Server, Analysis
Services, Qlikview',
 'Software Development, Business Intelligence, Big Data Analytics, Database
Administration, Data Science, Microsoft Azure, Spark, Cassandra, Object
Oriented',
 'Data Science, Node.js, Angularjs',
 'Data Science, Media Marketing, Resource Planning, Managed Services,
Display Advertising, Machine Learning, Python, Etl, Sql',
 'Data Analysis, Learning, Data Science, Computer Science, Communication
Skills',
 'Java, Hadoop, R, Machine Learning, Spark, Flume, Hdfs, Data Mining, Sas,
Big, Data Science, Cloudera, Impala, Bigdata',
 'Software Development, Core Java, Unit Testing, Customer Experience,
Problem Solving, Communication Skills, Mysql, Data Science, Sales
Management, Analytics']
```

In [83]:

```
company=[]
try:
    company_tags=driver.find_elements(By.XPATH,'//a[@class="ellipsis"]')
    for i in company_tags:
        company.append(i.text)
except NoSuchElementException:
```

```
    company.append('-')
```

In [84]:

```
company
```

Out[84]:

```
['Aakash Harit',
 'Data Science Network',
 'shravan Kumar Gaddam',
 'Shore Infotech India Pvt. Ltd',
 'MARSIAN Technologies LLP',
 'MARSIAN Technologies LLP',
 'Anik Agrawal',
 'Enerlytics Software Solutions Pvt Ltd',
 'subhas patel',
 'LibraryXProject',
 'Abhishek - Only Analytics Hiring - India and',
 'Apidel Technologies Division of Transpower',
 'Institute for Financial Management and Resear',
 'IFMR',
 'Balu Ramesh',
 'Techvantage Systems Pvt Ltd',
 'Asif Lucknowi',
 'Weupskill- Live Wire India',
 'InstaFinancials',
 'CBL Data Science Private Limited',
 'Kalpana Dumpala',
 'Innominds Software',
 'Mubarak',
 'MoneyTap',
 'Kushal Rastogi',
 'QuantMagnum Technologies Pvt. Ltd.',
 'Priyanka Akiri',
 'Infinitive Software Solutions',
 'Kapil Devang',
 'BISP Solutions',
 'Mahesh Babu Channa',
 'SocialPrachar.com',
 'Vaishnavi Kudalkar',
 'Codeachive learning',
 'Sakshi Chhikara',
 'BIZ INFOTECNO PRIVATE LIMITED',
 'Ruchi Dhote',
 'Bristlecone India Ltd',
 'Manisha Yadav',
 'Easi Tax',
 'Riya Rajesh',
 'Novelworx Digital Solutions',
 'Rashmi Bhattacharjee',
 'AXESTRACK SOFTWARE SOLUTIONS PRIVATE...',
 'Faizan Kareem',
```

'FirstTech Consaltants Pvt.Ltd',
'Rithika dadwal',
'Affine Analytics',
'Sandhya Khandagale',
'Compumatrice Multimedia Pvt Ltd',
'Shaun Rao',
'Exela Technologies',
'Deeparchi Sharma',
'ZIGRAM',
'Azahar Shaikh',
'NEAL ANALYTICS SERVICES PVT LTD',
'Manas',
'Autumn Leaf Consulting Services Private...',
'kumar',
'trainin',
'Sunil Vedula',
'Nanoprecise Sci Corp',
'Rajat Kumar',
'R.S Consultancy & Services',
'Dhruv Dev Dubey',
'NETAPS FOUNDATION',
'Avnish Mishra',
'RMS Risk Management Solutions',
'Jayanth N',
'Dollarbird Information Services Pvt, Ltd',
'Avodha',
'Nikitha Palaparthi',
'Priya Khare',
'Independent Consultant',
'Amit Sharma',
'ASCO consulting',
'Kanan',
'NY INST',
'Shashikant Chaudhary',
'3D India Staffing Research & Consulting...',
'Brad',
'O.C. Tanner',
'Rutuja Pawar',
'Demand Matrix',
'Madhusudhan Sridhar',
'MADHUSUDHAN SRIDHAR',
'Ankit Sinha',
'Suntech Global',
'Gaurav Chouhan',
'Strategic Consulting Lab',
'Rashi Kacker',
'Impel Labs Pvt. Ltd.',
'Ashwini',
'MRP Advisers',

```
 'Balaji Kolli',
 'Saras Solutions India Pvt Ltd',
 'Rajani Nagaraj',
 'WildJasmine',
 'ROHIT Kumar',
 'LNT Private Limited']
```

In [85]:

```
Company=company[1::2]
Company
```

Out[85]:

```
['Data Science Network',
 'Shore Infotech India Pvt. Ltd',
 'MARSIAN Technologies LLP',
 'Enerlytics Software Solutions Pvt Ltd',
 'LibraryXProject',
 'Apidel Technologies Division of Transpower',
 'IFMR',
 'Techvantage Systems Pvt Ltd',
 'Weupskill- Live Wire India',
 'CBL Data Science Private Limited',
 'Innominds Software',
 'MoneyTap',
 'QuantMagnum Technologies Pvt. Ltd.',
 'Infinitive Software Solutions',
 'BISP Solutions',
 'SocialPrachar.com',
 'Codeachive learning',
 'BIZ INFOTECNO PRIVATE LIMITED',
 'Bristlecone India Ltd',
 'Easi Tax',
 'Novelworx Digital Solutions',
 'AXESTRACK SOFTWARE SOLUTIONS PRIVATE...',
 'FirstTech Consaltants Pvt.Ltd',
 'Affine Analytics',
 'Compumatrice Multimedia Pvt Ltd',
 'Exela Technologies',
 'ZIGRAM',
 'NEAL ANALYTICS SERVICES PVT LTD',
 'Autumn Leaf Consulting Services Private...',
 'trainin',
 'Nanoprecise Sci Corp',
 'R.S Consultancy &amp;amp; Services',
 'NETAPS FOUNDATION',
 'RMS Risk Management Solutions',
 'Dollarbird Information Services Pvt, Ltd',
 'Nikitha Palaparthi',
 'Independent Consultant',
 'ASCO consulting',
 'NY INST',
```

```
 '3D India Staffing Research &amp;amp; Consulting...',
 'O.C. Tanner',
 'Demand Matrix',
 'MADHUSUDHAN SRIDHAR',
 'Suntech Global',
 'Strategic Consulting Lab',
 'Impel Labs Pvt. Ltd.',
 'MRP Advisers',
 'Saras Solutions India Pvt Ltd',
 'WildJasmine',
 'LNT Private Limited']
```

In [86]:

```
print(len(Name),len(Designation),len(Company),len(Skills),len(Location))
```

```
50 50 50 50 49
```

In [87]:

```
import pandas as pd
db=pd.DataFrame({'Name':Name,"Designation":Designation,"Company":Company,"Skills":Skills})
db
```

Out[87]:

	Name	Designation	Company	Skills
0	Aakash Harit	HR Manager	Data Science Network	Classic ASP Developer, Internet Marketing Prof...
1	shravan Kumar Gaddam	Company Recruiter	Shore Infotech India Pvt. Ltd	.Net, Java, Data Science, Linux Administration...
2	MARSIAN Technologies LLP	Company HR	MARSIAN Technologies LLP	Data Science, Artificial Intelligence, Machine...
3	Anik Agrawal	Company Recruiter	Enerlytics Software Solutions Pvt Ltd	Mean Stack, javascript, angularjs, mongodb, We...
4	subhas patel	Founder CEO	LibraryXProject	Hadoop, Spark, Digital Strategy, Data Architec...
5	Abhishek - Only Analytics Hiring - India and	Recruitment Lead Consultant	Apidel Technologies Division of Transpower	Analytics, Business Intelligence, Business Ana...
6	Institute for Financial Management and Resear	Programme Manager	IFMR	Data Science
7	Balu Ramesh	HR Administrator	Techvantage Systems Pvt Ltd	Machine Learning, algorithms, Go Getter, Compu...
8	Asif Lucknowi	Director	Weupskill- Live Wire India	Technical Training, Software Development, Pres...
9	InstaFinancials	Human Resource	CBL Data Science Private Limited	Software Development, It Sales, Account Manage...

	Name	Designation	Company	Skills
10	Kalpana Dumpala	Executive Hiring	Innominds Software	Qa, Ui/ux, Java Developer, Java Architect, C++...
11	Mubarak	Company HR	MoneyTap	Business Intelligence, Data Warehousing, Data ...
12	Kushal Rastogi	Company HR	QuantMagnum Technologies Pvt. Ltd.	Office Administration, Hr Administration, tele...
13	Priyanka Akiri	HR Manager	Infinitive Software Solutions	Oracle Dba, Data Science, Data Warehousing, ET...
14	Kapil Devang	HR Manager	BISP Solutions	Big Data, Hadoop, Data Analytics, Data Science
15	Mahesh Babu Channa	HR Team Lead	SocialPrachar.com	Social Media, digital media maketing, seo, smm...
16	Vaishnavi Kudalkar	HR Executive	Codeachive learning	Data Science, Python, Data Analytics
17	Sakshi Chhikara	Assistant Manager HR	BIZ INFOTECNO PRIVATE LIMITED	React.js, Data Science, Java, Front End, Busin...
18	Ruchi Dhote	Senior Executive Talent Acquisition	Bristlecone India Ltd	Qlikview, Qlik Sense, Microsoft Azure, Power B...
19	Manisha Yadav	HR Executive	Easi Tax	Telecalling, Client Interaction, Marketing, Re...
20	Riya Rajesh	Manager Talent Acquisition	Novelworx Digital Solutions	Data Science
21	Rashmi Bhattacharjee	HR Head	AXESTRACK SOFTWARE SOLUTIONS PRIVATE...	Corporate Sales, Software Development, Softwar...
22	Faizan Kareem	HR MANAGER	FirstTech Consaltants Pvt.Ltd	Data Analytics, Data Science, Machine Learning...
23	Rithika dadwal	HR Recruiter	Affine Analytics	Data Science, Machine Learning, Python, R, Dee...
24	Sandhya Khandagale	HR Recruiter	Compumatrice Multimedia Pvt Ltd	Big Data, Data Science, Artificial Intelligenc...
25	Shaun Rao	Manager Human Resources	Exela Technologies	Java, Net, Angularjs, Hr, Infrastructure, Mana...
26	Deeparchi Sharma	Company Recruiter	ZIGRAM	Research, Digital Marketing, Analytics, Softwa...
27	Azahar Shaikh	Company Recruiter	NEAL ANALYTICS SERVICES PVT LTD	Data Science, Artificial Intelligence, Machine...

	Name	Designation	Company	Skills
28	Manas	Lead Talent acquisition	Autumn Leaf Consulting Services Private...	Software Architecture, Vp Engineering, Product...
29	kumar	Proprietor	trainin	Data Science, Hadoop, Rpas, Devops, Python, Aw...
30	Sunil Vedula	CEO	Nanoprecise Sci Corp	Signal Processing, Machine Learning, Neural Ne...
31	Rajat Kumar	Founder CEO	R.S Consultancy & Services	Web Technologies, Project Management, Software...
32	Dhruv Dev Dubey	Company Recruitment Head	NETAPS FOUNDATION	Server Administartion, Verilog, Vhdl, Digital ...
33	Avnish Mishra	Senior Technical & Corporate Recruiter	RMS Risk Management Solutions	Data Science, Machine Learning, Deep Learning,...
34	Jayanth N	Project Manager	Dollarbird Information Services Pvt, Ltd	Data Analytics, Managed Services, Team Leading...
35	Avodha	Business Development Associate	Nikitha Palaparthi	Ethical Hacking, Security Operations Center, S...
36	Priya Khare	Senior Manager	Independent Consultant	Data Science, Artificial Intelligence, analyti...
37	Amit Sharma	Consultant	ASCO consulting	Machine Learning, Artificial Intelligence, Dat...
38	Kanan	senior technology instructor	NY INST	C, C++, Artificial Intelligence, Python, Php, ...
39	Shashikant Chaudhary	HR Recruiter/HR Excutive	3D India Staffing Research & Consulting...	Relationship Management, Retail Sales, Private...
40	Brad	Manager, Technical Recruiting	O.C. Tanner	Data Science, Software Engineering
41	Rutuja Pawar	Technical Recruiter	Demand Matrix	Data Science, Big Data Analytics, Digital Mark...
42	Madhusudhan Sridhar	Erp Implementer	MADHUSUDHAN SRIDHAR	Data Science, Recruitment, Salary
43	Ankit Sinha	Head Analytics	Suntech Global	B.Tech, Tableau, Statistics, R, Analytics, Tim...
44	Gaurav Chouhan	Chief Technical Officer	Strategic Consulting Lab	Software Development, Business Intelligence, B...
45	Rashi Kacker	Sr Product Manager	Impel Labs Pvt. Ltd.	Data Science, Node.js, Angularjs

	Name	Designation	Company	Skills
46	Ashwini	Director Global Delivery	MRP Advisers	Data Science, Media Marketing, Resource Planni...
47	Balaji Kolli	Co Founder	Saras Solutions India Pvt Ltd	Data Analysis, Learning, Data Science, Compute...
48	Rajani Nagaraj	HR Manager	WildJasmine	Java, Hadoop, R, Machine Learning, Spark, Flum...
49	ROHIT Kumar	Architect	LNT Private Limited	Software Development, Core Java, Unit Testing,...

In [ ]:
```
Location is not availabe for all 50 data sets.
```

In [ ]:
```
Answer 3
```

In [88]:
```
import selenium
import pandas as pd
import time
from bs4 import BeautifulSoup
from selenium import webdriver
from selenium.common.exceptions import StaleElementReferenceException,
NoSuchElementException
import requests
from selenium.webdriver.common.by import By
```

In [89]:
```
driver=webdriver.Chrome(r"C:\Users\zarna\Downloads\chromedriver_win32
(1)\chromedriver.exe")
```

```
C:\Users\zarna\AppData\Local\Temp\ipykernel_169148\1240552591.py:1:
DeprecationWarning: executable_path has been deprecated, please pass in a
Service object
  driver=webdriver.Chrome(r"C:\Users\zarna\Downloads\chromedriver_win32
(1)\chromedriver.exe")
```

In [90]:
```
driver.get("https://www.statisticstimes.com/")
```

In [94]:
```
click=driver.find_element(By.CLASS_NAME,'ec')
click.click()
```

In [95]:
```
Rank=[]
State=[]
GSDP1819=[]
GSDP1920=[]
Share1819=[]
GDP=[]
```

In [96]:
```
try:
    rank_tags=driver.find_elements(By.XPATH,'//td[@class="data1"]')
    for i in rank_tags[0:34]:
```

```
        Rank.append(i.text)
except NoSuchElementException:
    Rank.append('-')
```

```
Rank
```

Out[96]:

```
['1',
 '2',
 '3',
 '4',
 '5',
 '6',
 '7',
 '8',
 '9',
 '10',
 '11',
 '12',
 '13',
 '14',
 '15',
 '16',
 '17',
 '18',
 '19',
 '20',
 '21',
 '22',
 '23',
 '24',
 '25',
 '26',
 '27',
 '28',
 '29',
 '30',
 '31',
 '32',
 '33',
 '']
```

In [98]:

```
rank=[]
rank=Rank[0:33]
rank
```

Out[98]:

```
['1',
 '2',
 '3',
 '4',
```

```
 '5',
 '6',
 '7',
 '8',
 '9',
 '10',
 '11',
 '12',
 '13',
 '14',
 '15',
 '16',
 '17',
 '18',
 '19',
 '20',
 '21',
 '22',
 '23',
 '24',
 '25',
 '26',
 '27',
 '28',
 '29',
 '30',
 '31',
 '32',
 '33']
```

In [99]:

```
try:
    State_tags=driver.find_elements(By.XPATH,'//td[@class="name"]')
    for i in State_tags[0:33]:
        State.append(i.text)
except NoSuchElementException:
    State.append('-')

try:
    GSDP1819_tags=driver.find_elements(By.XPATH,'//td[@class="data
sorting_1"]')
    for i in GSDP1819_tags[0:33]:
        GSDP1819.append(i.text)
except NoSuchElementException:
    GSDP1819.append('-')
```

In [ ]:

In [100]:

```
data=[]
try:
```

```
    data_tags=driver.find_elements(By.XPATH,'//td[@class="data"]')
    for i in data_tags[0:164]:
        data.append(i.text)
except NoSuchElementException:
    data.append('-')
```

```
data
```

Out[100]:

```
['-',
 '13.94%',
 '399.921',
 '-',
 '2,039,074',
 '1,845,853',
 '8.63%',
 '247.629',
 '1,312,929',
 '1,215,307',
 '1,687,818',
 '8.39%',
 '240.726',
 '1,166,817',
 '1,123,982',
 '-',
 '7.96%',
 '228.290',
 '-',
 '1,186,379',
 '1,631,977',
 '7.91%',
 '226.806',
 '1,156,039',
 '1,091,077',
 '1,253,832',
 '5.77%',
 '165.556',
 '793,223',
 '739,525',
 '1,020,989',
 '4.99%',
 '143.179',
 '711,627',
 '677,428',
 '972,782',
 '4.57%',
 '131.083',
 '672,018',
 '621,301',
 '969,604',
```

'4.56%',
'130.791',
'663,258',
'612,828',
'906,672',
'4.29%',
'122.977',
'561,801',
'522,009',
'-',
'4.14%',
'118.733',
'-',
'559,412',
'856,112',
'4.10%',
'117.703',
'634,408',
'590,569',
'831,610',
'3.89%',
'111.519',
'572,240',
'531,085',
'611,804',
'2.81%',
'80.562',
'414,977',
'375,651',
'574,760',
'2.79%',
'79.957',
'418,868',
'397,669',
'521,275',
'2.58%',
'74.098',
'396,499',
'376,877',
'-',
'1.67%',
'47.982',
'-',
'234,048',
'329,180',
'1.61%',
'46.187',
'243,477',
'231,182',

'328,598',
'1.57%',
'45.145',
'240,036',
'224,986',
'-',
'1.30%',
'37.351',
'-',
'193,273',
'-',
'0.83%',
'23.690',
'-',
'112,755',
'165,472',
'0.81%',
'23.369',
'124,403',
'117,851',
'80,449',
'0.39%',
'11.115',
'63,408',
'57,787',
'55,984',
'0.26%',
'7.571',
'40,583',
'36,963',
'-',
'0.22%',
'6.397',
'-',
'31,192',
'38,253',
'0.18%',
'5.230',
'25,093',
'23,013',
'36,572',
'0.18%',
'5.086',
'26,695',
'24,682',
'32,496',
'0.15%',
'4.363',
'20,017',

```
 '18,722',
 '31,790',
 '0.15%',
 '4.233',
 '20,673',
 '19,300',
 '-',
 '0.14%',
 '4.144',
 '-',
 '17,647',
 '-',
 '0.13%',
 '3.737',
 '-',
 '16,676',
 '26,503',
 '0.12%',
 '3.385',
 '18,797',
 '16,478',
 '-',
 '-',
 '-',
 '-']
```

In [101]:

```
GSDP1920=data[0::5]
GSDP1920
```

Out[101]:

```
['-',
 '1,845,853',
 '1,687,818',
 '-',
 '1,631,977',
 '1,253,832',
 '1,020,989',
 '972,782',
 '969,604',
 '906,672',
 '-',
 '856,112',
 '831,610',
 '611,804',
 '574,760',
 '521,275',
 '-',
 '329,180',
 '328,598',
 '-',
```

```
 '-',
 '165,472',
 '80,449',
 '55,984',
 '-',
 '38,253',
 '36,572',
 '32,496',
 '31,790',
 '-',
 '-',
 '26,503',
 '-']
```

In [102]:

```
Share1819=data[1::5]
Share1819
```

Out[102]:

```
['13.94%',
 '8.63%',
 '8.39%',
 '7.96%',
 '7.91%',
 '5.77%',
 '4.99%',
 '4.57%',
 '4.56%',
 '4.29%',
 '4.14%',
 '4.10%',
 '3.89%',
 '2.81%',
 '2.79%',
 '2.58%',
 '1.67%',
 '1.61%',
 '1.57%',
 '1.30%',
 '0.83%',
 '0.81%',
 '0.39%',
 '0.26%',
 '0.22%',
 '0.18%',
 '0.18%',
 '0.15%',
 '0.15%',
 '0.14%',
 '0.13%',
 '0.12%',
```

```
 '-']
```

In [103]:

```
GDP=data[2::5]
GDP
```

Out[103]:

```
['399.921',
 '247.629',
 '240.726',
 '228.290',
 '226.806',
 '165.556',
 '143.179',
 '131.083',
 '130.791',
 '122.977',
 '118.733',
 '117.703',
 '111.519',
 '80.562',
 '79.957',
 '74.098',
 '47.982',
 '46.187',
 '45.145',
 '37.351',
 '23.690',
 '23.369',
 '11.115',
 '7.571',
 '6.397',
 '5.230',
 '5.086',
 '4.363',
 '4.233',
 '4.144',
 '3.737',
 '3.385',
 '-']
```

In [107]:

```
print(len(State),len(GSDP1819),len(GSDP1920),len(Share1819),len(GDP),len(ra
nk))
```

```
33 33 33 33 33 33
```

In [108]:

```
import pandas as pd
db=pd.DataFrame({"Rank":rank,"State":State,"GSDP 18-19":GSDP1819,"GSDP 19-
20":GSDP1920,"Share 18-19":Share1819,"GDP":GDP})
db
```

Out[108]:

	Rank	State	GSDP 18-19	GSDP 19-20	Share 18-19	GDP
0	1	Maharashtra	2,632,792	-	13.94%	399.921
1	2	Tamil Nadu	1,630,208	1,845,853	8.63%	247.629
2	3	Uttar Pradesh	1,584,764	1,687,818	8.39%	240.726
3	4	Gujarat	1,502,899	-	7.96%	228.290
4	5	Karnataka	1,493,127	1,631,977	7.91%	226.806
5	6	West Bengal	1,089,898	1,253,832	5.77%	165.556
6	7	Rajasthan	942,586	1,020,989	4.99%	143.179
7	8	Andhra Pradesh	862,957	972,782	4.57%	131.083
8	9	Telangana	861,031	969,604	4.56%	130.791
9	10	Madhya Pradesh	809,592	906,672	4.29%	122.977
10	11	Kerala	781,653	-	4.14%	118.733
11	12	Delhi	774,870	856,112	4.10%	117.703
12	13	Haryana	734,163	831,610	3.89%	111.519
13	14	Bihar	530,363	611,804	2.81%	80.562
14	15	Punjab	526,376	574,760	2.79%	79.957
15	16	Odisha	487,805	521,275	2.58%	74.098
16	17	Assam	315,881	-	1.67%	47.982
17	18	Chhattisgarh	304,063	329,180	1.61%	46.187
18	19	Jharkhand	297,204	328,598	1.57%	45.145
19	20	Uttarakhand	245,895	-	1.30%	37.351
20	21	Jammu & Kashmir	155,956	-	0.83%	23.690
21	22	Himachal Pradesh	153,845	165,472	0.81%	23.369
22	23	Goa	73,170	80,449	0.39%	11.115
23	24	Tripura	49,845	55,984	0.26%	7.571
24	25	Chandigarh	42,114	-	0.22%	6.397
25	26	Puducherry	34,433	38,253	0.18%	5.230
26	27	Meghalaya	33,481	36,572	0.18%	5.086
27	28	Sikkim	28,723	32,496	0.15%	4.363
28	29	Manipur	27,870	31,790	0.15%	4.233
29	30	Nagaland	27,283	-	0.14%	4.144
30	31	Arunachal Pradesh	24,603	-	0.13%	3.737
31	32	Mizoram	22,287	26,503	0.12%	3.385
32	33	Andaman & Nicobar Islands	-	-	-	-

In [ ]:

```
Answer 4
```

In [115]:

```
import selenium
import pandas as pd
import time
from bs4 import BeautifulSoup
from selenium import webdriver
```

```
from selenium.common.exceptions import StaleElementReferenceException,
NoSuchElementException
import requests
from selenium.webdriver.common.by import By
```

In [116]:

```
driver=webdriver.Chrome(r"C:\Users\zarna\Downloads\chromedriver_win32
(1)\chromedriver.exe")
```

```
C:\Users\zarna\AppData\Local\Temp\ipykernel_169148\1240552591.py:1:
DeprecationWarning: executable_path has been deprecated, please pass in a
Service object
  driver=webdriver.Chrome(r"C:\Users\zarna\Downloads\chromedriver_win32
(1)\chromedriver.exe")
```

In [117]:

```
driver.get("https://github.com/")
```

In [ ]:

```
Not able to go on trending option
```

In [118]:

```
Repository_title=[]
Repository_description=[]
Contributors_count=[]
Language_used=[]
```

In [119]:

```
try:
    title_tags=driver.find_elements(By.XPATH,'//h2[@class="h3 lh-
condensed"]')
    for i in title_tags:
        Repository_title.append(i.text)
except NoSuchElementException:
    Repository_title.append("-")

Repository_title
```

Out[119]:

```
['Ryujinx / Ryujinx',
 'gventuri / pandas-ai',
 'mlc-ai / mlc-llm',
 'xtekky / gpt4free',
 'AIGC-Audio / AudioGPT',
 'pixiv / ChatVRM',
 'Lissy93 / dashy',
 'karpathy / nanoGPT',
 'mishalhossin / Discord-Chatbot-Gpt4Free',
 'immich-app / immich',
 'memorysafety / sudo-rs',
 'awesome-selfhosted / awesome-selfhosted',
 'xtekky / chatgpt-clone',
 'amigoscode / full-stack-professional',
 'georgia-tech-db / eva',
 'Azure / Azure-Sentinel',
 'IntelligenzaArtificiale / Free-AUTO-GPT-with-NO-API',
 'nextcloud / server',
```

```
 'google-research / google-research',
 'drizzle-team / drizzle-orm',
 'ripienaar / free-for-dev',
 'frappe / frappe',
 'mui / material-ui',
 'jesseduffield / lazygit',
 'vlang / v']
```

In [120]:

```
try:
    description_tags=driver.find_elements(By.XPATH,'//p[@class="col-9
color-fg-muted my-1 pr-4"]')
    for i in description_tags:
        Repository_description.append(i.text)
except NoSuchElementException:
    Repository_description.append('-')

Repository_description
```

Out[120]:

```
['Experimental Nintendo Switch Emulator written in C#',
 'Pandas AI is a Python library that integrates generative artificial
intelligence capabilities into Pandas, making dataframes conversational',
 "Enable everyone to develop, optimize and deploy AI models natively on
everyone's devices.",
 "decentralising the Ai Industry, just some language model api's...",
 'AudioGPT: Understanding and Generating Speech, Music, Sound, and Talking
Head',
 '🚀 A self-hostable personal dashboard built for you. Includes status-
checking, widgets, themes, icon packs, a UI editor and tons more!',
 'The simplest, fastest repository for training/finetuning medium-sized
GPTs.',
 'This is a Python-based Discord Chatbot. This is all free due to the
GPT4FREE project',
 'Self-hosted photo and video backup solution directly from your mobile
phone.',
 'A memory safe implementation of sudo and su.',
 'A list of Free Software network services and web applications which can
be hosted on your own servers',
 'ChatGPT interface with better UI',
 'AI-Relational Database System | SQL meets Deep Learning',
 'Cloud-native SIEM for intelligent security analytics for your entire
enterprise.',
 'Free AUTOGPT with NO API is a repository that offers a simple version of
Autogpt, an autonomous AI agent capable of performing tasks independently.
Unlike other versions, our implementation does not rely on any paid OpenAI
API, making it accessible to anyone.',
 '☁ Nextcloud server, a safe home for all your data',
 'Google Research',
 'TypeScript ORM for SQL',
```

```
 'A list of SaaS, PaaS and IaaS offerings that have free tiers of interest
to devops and infradev',
 'Low code web framework for real world applications, in Python and
Javascript',
 "MUI Core: Ready-to-use foundational React components, free forever. It
includes Material UI, which implements Google's Material Design.",
 'simple terminal UI for git commands',
 'Simple, fast, safe, compiled language for developing maintainable
software. Compiles itself in <1s with zero library dependencies. Supports
automatic C => V translation. https://vlang.io']
```

In [121]:

```
try:
    count_tags=driver.find_elements(By.XPATH,'//span[@class="d-inline-block
ml-0 mr-3"]')
    for i in count_tags:
        Contributors_count.append(i.text)
except NoSuchElementException:
    Contributors_count.append('-')

Contributors_count
```

Out[121]:

```
['C#',
 'Python',
 'Python',
 'Python',
 'Python',
 'TypeScript',
 'Vue',
 'Python',
 'Python',
 'Dart',
 'Rust',
 'Makefile',
 'Python',
 'Java',
 'Python',
 'Jupyter Notebook',
 'Jupyter Notebook',
 'PHP',
 'Jupyter Notebook',
 'TypeScript',
 'HTML',
 'Python',
 'JavaScript',
 'Go',
 'V']
```

In [124]:

```
Data=[]
```

```
try:
    data_tags=driver.find_elements(By.XPATH,'//a[@class="Link--muted d-
inline-block mr-3"]')
    for i in data_tags:
        Data.append(i.text)
except NoSuchElementException:
    Data.append('-')
```

```
Data
```

Out[124]:

```
['19,526',
 '1,962',
 '2,885',
 '163',
 '4,747',
 '286',
 '28,826',
 '6,603',
 '6,996',
 '592',
 '292',
 '42',
 '11,180',
 '808',
 '19,532',
 '2,273',
 '280',
 '95',
 '10,684',
 '383',
 '1,533',
 '37',
 '129,988',
 '7,502',
 '1,853',
 '580',
 '226',
 '55',
 '1,388',
 '135',
 '3,302',
 '2,298',
 '959',
 '109',
 '22,314',
 '3,491',
 '29,008',
 '7,165',
 '5,171',
```

```
 '108',
 '69,581',
 '7,573',
 '4,750',
 '2,470',
 '86,391',
 '29,751',
 '34,571',
 '1,399',
 '33,551',
 '2,039']
```

In [125]:

```
Language_used=Data[1::2]
```

In [126]:

```
Language_used
```

Out[126]:

```
['1,962',
 '163',
 '286',
 '6,603',
 '592',
 '42',
 '808',
 '2,273',
 '95',
 '383',
 '37',
 '7,502',
 '580',
 '55',
 '135',
 '2,298',
 '109',
 '3,491',
 '7,165',
 '108',
 '7,573',
 '2,470',
 '29,751',
 '1,399',
 '2,039']
```

In [127]:

```
print(len(Repository_title),len(Repository_description),len(Contributors_co
unt),len(Language_used))
25 23 25 25
```

In [128]:

```
import pandas as pd
db=pd.DataFrame({"Repository title":Repository_title,"Contributors
count":Language_used,"Language used":Contributors_count})
db
```

Out[128]:

	Repository title	Contributors count	Language used
0	Ryujinx / Ryujinx	1,962	C#
1	gventuri / pandas-ai	163	Python
2	mlc-ai / mlc-llm	286	Python
3	xtekky / gpt4free	6,603	Python
4	AIGC-Audio / AudioGPT	592	Python
5	pixiv / ChatVRM	42	TypeScript
6	Lissy93 / dashy	808	Vue
7	karpathy / nanoGPT	2,273	Python
8	mishalhossin / Discord-Chatbot-Gpt4Free	95	Python
9	immich-app / immich	383	Dart
10	memorysafety / sudo-rs	37	Rust
11	awesome-selfhosted / awesome-selfhosted	7,502	Makefile
12	xtekky / chatgpt-clone	580	Python
13	amigoscode / full-stack-professional	55	Java
14	georgia-tech-db / eva	135	Python
15	Azure / Azure-Sentinel	2,298	Jupyter Notebook
16	IntelligenzaArtificiale / Free-AUTO-GPT-with-N...	109	Jupyter Notebook
17	nextcloud / server	3,491	PHP
18	google-research / google-research	7,165	Jupyter Notebook
19	drizzle-team / drizzle-orm	108	TypeScript
20	ripienaar / free-for-dev	7,573	HTML
21	frappe / frappe	2,470	Python
22	mui / material-ui	29,751	JavaScript
23	jesseduffield / lazygit	1,399	Go
24	vlang / v	2,039	V

In [ ]:

In [ ]:

In [ ]:

In [ ]:

In [111]:

In [ ]:

In [112]:

In [ ]:

In [ ]:

In [ ]:

In [ ]:

In [ ]:

In [ ]:

In [ ]:

In [ ]:

In [ ]:

In [ ]:

In [ ]:

In [ ]:

In [ ]:

In [ ]:

In [ ]:

In [ ]:

In [ ]:

In [ ]:

In [ ]:

In [ ]:

In [ ]:

In [ ]:

In [ ]:

In [ ]:

In [ ]:

In [ ]:

In [ ]:

In [ ]:

In [ ]:

In [ ]:

In [ ]:

In [ ]:

In [ ]:

In [ ]:

In [ ]:

In [ ]:

In [ ]:

In [5]:

In [ ]:

In [4]:

In [6]:

In [ ]:

In [ ]:

In [ ]: