Taller 8

Métodos Computacionales para Políticas Públicas - URosario

Entrega: viernes 6-nov-2020 11:59 PM

[Nicolás Garcés R]

[nicolas.garces@urosario.edu.co]

Instrucciones:

- Guarde una copia de este *Jupyter Notebook* en su computador, idealmente en una carpeta destinada al material del curso.
- Modifique el nombre del archivo del *notebook*, agregando al final un guión inferior y su nombre y apellido, separados estos últimos por otro guión inferior. Por ejemplo, mi *notebook* se llamaría: mcpp_taller8_santiago_matallana
- Marque el *notebook* con su nombre y e-mail en el bloque verde arriba. Reemplace el texto "[Su nombre acá]" con su nombre y apellido. Similar para su e-mail.
- Desarrolle la totalidad del taller sobre este *notebook*, insertando las celdas que sea necesario debajo de cada pregunta. Haga buen uso de las celdas para código y de las celdas tipo *markdown* según el caso.
- Recuerde salvar periódicamente sus avances.
- Cuando termine el taller:
    1. Descárguelo en PDF. Si tiene algún problema con la conversión, descárguelo en HTML.
    2. Suba todos los archivos a su repositorio en GitHub, en una carpeta destinada exclusivamente para este taller, antes de la fecha y hora límites.

1. [1 punto]

Usando expresiones regulares extraiga en una lista todos los números presentes en el siguiente objeto de Python:

ob1 = "JEFF BEZOS, the founder of Amazon, has reached a divorce settlement with his wife, MacKenzie. Mr Bezos will keep all the shares in the Washington Post and Blue Origin, a space-exploration firm, as well as 75% of the couple's Amazon stock. Mrs Bezos will retain a 4% stake in the tech giant, worth nearly $36bn, which is likely to make her the third-richest woman alive when the divorce is finalised."

In [7]:
```
ob1 = "JEFF BEZOS, the founder of Amazon, has reached a divorce settlem
ent with his wife, MacKenzie. Mr Bezos will keep all the shares in the
 Washington Post and Blue Origin, a space-exploration firm, as well as
 75% of the couple's Amazon stock. Mrs Bezos will retain a 4% stake in
 the tech giant, worth nearly $36bn, which is likely to make her the th
ird-richest woman alive when the divorce is finalised."
```

In [8]:
```
import re
re.findall("[0-9]+",ob1)
```

Out[8]:
```
['75', '4', '36']
```

2. [1 punto]

Usando expresiones regulares ahora extraiga de *ob1* sólo los números que correspondan a porcentajes.

In [12]:
```
re.findall("([0-9]+)%",ob1)
```

Out[12]: ['75', '4']

3. [2 puntos]

Usando expresiones regulares, escriba una función de Python que reciba una fecha en formato **Marzo 7, 2019** y retorne la fecha en formato **2019-07-03**

In [1]:
```
def dater(x):
    import re
    y = re.findall("([A-Z][a-zA-Z]*)",x)
    if y[0] =="Enero":
        yy= "01"
    elif y[0] =="Febrero":
        yy= "02"
    elif y[0] =="Marzo":
        yy= "03"
    elif y[0] =="Abril":
        yy= "04"
    elif y[0] =="Mayo":
        yy= "05"
    elif y[0] =="Junio":
        yy= "06"
    elif y[0] =="Julio":
        yy= "07"
    elif y[0] =="Agosto":
        yy= "08"
    elif y[0] =="Septiembre":
        yy= "09"
    elif y[0] =="Octubre":
        yy= "10"
    elif y[0] =="Noviembre":
        yy= "11"
    elif y[0] =="Diciembre":
        yy= "12"
    else:
        yy="00"
    z = re.findall("[0-9]+",x)
```

```
    zz= z[0]
    if len(zz) <= 1:
        zz= "0"+zz
    else:
        zz = zz
    zzz= z[1]
    re= zzz + "-" + zz + "-"+ yy
    return re
```

In [3]:
```
x = "Septiembre 25, 2019"
dater(x)
```

Out[3]: '2019-25-09'

In [4]:
```
x = "Marzo 7, 2019"
dater(x)
```

Out[4]: '2019-07-03'

4. [3 puntos]

ob2 es un string que reune una lista de clases en una universidad. Use expresiones regulares para extraer los códigos de cada una de las clases. Ejemplo: El código de la clase **COMPSCI 143 (Spring 2012): Machine Learning** es 143.

ob2 = "COMPSCI 270 (Spring 2019): Introduction to Artificial Intelligence. COMPSCI 590.2 (Fall 2018): Computational Microeconomics: Game Theory, Social Choice, and Mechanism Design. COMPSCI 223 (Spring 2018): Computational Microeconomics. COMPSCI 570 (Fall 2017): Artificial Intelligence. COMPSCI 590.3 (Fall 2017) / 590.1 (Spring 2018): Ethics and AI. COMPSCI 590.2 (Spring 2017): Computation, Information, and Learning in Market Design. COMPSCI 590.4 (Spring 2016): Computational Microeconomics: Game Theory, Social Choice, and Mechanism Design. COMPSCI 290.4/590.4 (Spring 2015): Crowdsourcing Societal Tradeoffs. COMPSCI 570 (Fall 2014): Artificial Intelligence. COMPSCI 590.4 (Spring 2014): Computational Microeconomics: Game Theory, Social Choice, and Mechanism Design. COMPSCI 590.1 (Fall 2012): Linear and Integer Programming. COMPSCI 173 (Spring 2012):

Computational Microeconomics. COMPSCI 296.1 (Fall 2011): Computational Microeconomics: Game Theory, Social Choice, and Mechanism Design. COMPSCI 296.1 (Fall 2010): Linear and Integer Programming. COMPSCI 173 (Spring 2010): Computational Microeconomics. COMPSCI 196.1/296.1 (Fall 2009): Computational Microeconomics: Game Theory, Social Choice, and Mechanism Design. COMPSCI 170 (Spring 2009): Introduction to Artificial Intelligence. COMPSCI 270 (Fall 2008): Artificial Intelligence. COMPSCI 196/296.2 (Spring 2008): Linear and Integer Programming. COMPSCI 196.2 (Fall 2007): Introduction to Computational Economics. COMPSCI 296.3 (Spring 2007): Topics in Computational Economics. COMPSCI 296.2 (Fall 2006): Computational Game Theory and Mechanism Design."

In [24]:
```
ob2 = "COMPSCI 270 (Spring 2019): Introduction to Artificial Intelligen
ce. COMPSCI 590.2 (Fall 2018): Computational Microeconomics: Game Theor
y, Social Choice, and Mechanism Design. COMPSCI 223 (Spring 2018): Comp
utational Microeconomics. COMPSCI 570 (Fall 2017): Artificial Intellige
nce. COMPSCI 590.3 (Fall 2017) / 590.1 (Spring 2018): Ethics and AI. CO
MPSCI 590.2 (Spring 2017): Computation, Information, and Learning in Ma
rket Design. COMPSCI 590.4 (Spring 2016): Computational Microeconomics:
 Game Theory, Social Choice, and Mechanism Design. COMPSCI 290.4/590.4
 (Spring 2015): Crowdsourcing Societal Tradeoffs. COMPSCI 570 (Fall 201
4): Artificial Intelligence. COMPSCI 590.4 (Spring 2014): Computational
 Microeconomics: Game Theory, Social Choice, and Mechanism Design. COMP
SCI 590.1 (Fall 2012): Linear and Integer Programming. COMPSCI 173 (Spr
ing 2012): Computational Microeconomics. COMPSCI 296.1 (Fall 2011): Com
putational Microeconomics: Game Theory, Social Choice, and Mechanism De
sign. COMPSCI 296.1 (Fall 2010): Linear and Integer Programming. COMPSC
I 173 (Spring 2010): Computational Microeconomics. COMPSCI 196.1/296.1
 (Fall 2009): Computational Microeconomics: Game Theory, Social Choice,
 and Mechanism Design. COMPSCI 170 (Spring 2009): Introduction to Artif
icial Intelligence. COMPSCI 270 (Fall 2008): Artificial Intelligence. C
OMPSCI 196/296.2 (Spring 2008): Linear and Integer Programming. COMPSCI
 196.2 (Fall 2007): Introduction to Computational Economics. COMPSCI 29
6.3 (Spring 2007): Topics in Computational Economics. COMPSCI 296.2 (Fa
ll 2006): Computational Game Theory and Mechanism Design."
```

In [26]:
```
import re
y = re.findall("[A-Z] ([0-9.]*) ",ob2)
x = re.findall("([0-9.]+)/",ob2)
```

```
z = re.findall("/([0-9.]+)",ob2)
final = x + y + z
final
```

Out[26]:
```
['290.4',
 '196.1',
 '196',
 '270',
 '590.2',
 '223',
 '570',
 '590.3',
 '590.2',
 '590.4',
 '570',
 '590.4',
 '590.1',
 '173',
 '296.1',
 '296.1',
 '173',
 '170',
 '270',
 '196.2',
 '296.3',
 '296.2',
 '590.4',
 '296.1',
 '296.2']
```

5. [5 puntos]

ob3 es un string que reune una lista de publicaciones. Use expresiones regulares para extraer todos los *Journals* en los cuales el autor ha publicado. Ejemplo: El paper **Bail, CA. "The configuration of symbolic boundaries against immigrants in Europe." American Sociological Review 73.1 (January 1, 2008): 37-59. Full Text** fue publicado en el Journal *American Sociological Review*

ob3 = "Bail, CA, Argyle, LP, Brown, TW, Bumpus, JP, Chen, H, Hunzaker, MBF, Lee, J, Mann, M, Merhout, F, and Volfovsky, A. \"Exposure to opposing views on social media can increase political polarization.\" Proceedings of the National Academy of Sciences of the United States of America 115.37 (September 2018): 9216-9221. Full Text Open Access Copy.\n", "Bail, CA, Merhout, F, and Ding, P. \"Using Internet search data to examine the relationship between anti-Muslim and pro-ISIS sentiment in U.S. counties.\" Science Advances 4.6 (June 6, 2018): eaao5948-null. Full Text Open Access Copy.\n", "Bail, CA, Brown, TW, and Mann, M. \"Channeling Hearts and Minds: Advocacy Organizations, Cognitive-Emotional Currents, and Public Conversation.\" American Sociological Review 82.6 (December 1, 2017): 1188-1213. Full Text.\n", "Bail, CA. \"Taming Big Data: Using App Technology to Study Organizational Behavior on Social Media.\" Sociological Methods and Research 46.2 (March 1, 2017): 189-217. Full Text.\n", "McDonnell, TE, Bail, CA, and Tavory, I. \"A Theory of Resonance.\" Sociological Theory 35.1 (March 1, 2017): 1-14. Full Text.\n", "Bail, CA. \"Combining natural language processing and network analysis to examine how advocacy organizations stimulate conversation on social media.\" Proceedings of the National Academy of Sciences of the United States of America 113.42 (October 2016): 11823-11828. Full Text.\n", "Bail, CA. \"Emotional Feedback and the Viral Spread of Social Media Messages About Autism Spectrum Disorders.\" American journal of public health 106.7 (July 2016): 1173-1180. Full Text.\n", "Bail, CA. \"The public life of secrets: Deception, disclosure, and discursive framing in the policy process.\" Sociological Theory 33.2 (January 1, 2015): 97-124. Full Text.\n", "Bail, CA. \"The cultural environment: Measuring culture with big data.\" Theory and Society 43.3 (January 1, 2014): 465-524. Full Text.\""

In [6]:
```
ob3 = '''"Bail, CA, Argyle, LP, Brown, TW, Bumpus, JP, Chen, H, Hunzaker, MBF, Lee, J, Mann, M, Merhout, F, and Volfovsky, A. "Exposure to opposing views on social media can increase political polarization." Proceedings of the National Academy of Sciences of the United States of America 115.37 (September 2018): 9216-9221. Full Text Open Access Copy.\n",
 "Bail, CA, Merhout, F, and Ding, P. "Using Internet search data to examine the relationship between anti-Muslim and pro-ISIS sentiment in U.S. counties." Science Advances 4.6 (June 6, 2018): eaao5948-null. Full Text Open Access Copy.\n", "Bail, CA, Brown, TW, and Mann, M. "Channeling Hearts and Minds: Advocacy Organizations, Cognitive-Emotional Currents, and Public Conversation." American Sociological Review 82.6 (December 1, 2017): 1188-1213. Full Text.\n", "Bail, CA. "Taming Big Data: Using App Technology to Study Organizational Behavior on Social Media." S
```

```
ociological Methods and Research 46.2 (March 1, 2017): 189-217. Full Te
xt.\n", "McDonnell, TE, Bail, CA, and Tavory, I. "A Theory of Resonanc
e." Sociological Theory 35.1 (March 1, 2017): 1-14. Full Text.\n", "Bai
l, CA. "Combining natural language processing and network analysis to e
xamine how advocacy organizations stimulate conversation on social medi
a." Proceedings of the National Academy of Sciences of the United State
s of America 113.42 (October 2016): 11823-11828. Full Text.\n", "Bail,
 CA. "Emotional Feedback and the Viral Spread of Social Media Messages
 About Autism Spectrum Disorders." American journal of public health 10
6.7 (July 2016): 1173-1180. Full Text.\n", "Bail, CA. "The public life
 of secrets: Deception, disclosure, and discursive framing in the polic
y process." Sociological Theory 33.2 (January 1, 2015): 97-124. Full Te
xt.\n", "Bail, CA. "The cultural environment: Measuring culture with bi
g data." Theory and Society 43.3 (January 1, 2014): 465-524. Full Tex
t."""'
```

In [7]:
```
re.findall('''." ([A-Z \s a-z]+) [0-9]*''',ob3)
```

Out[7]:
```
['Proceedings of the National Academy of Sciences of the United States
of America',
 'Science Advances',
 'American Sociological Review',
 'Sociological Methods and Research',
 'Sociological Theory',
 'Proceedings of the National Academy of Sciences of the United States
of America',
 'American journal of public health',
 'Sociological Theory',
 'Theory and Society']
```

6. [10 puntos]

Vamos a hacer "scraping" a esta página: https://archive.ics.uci.edu/ml/datasets.php, que contiene un listado de 559 bases de datos que hacen parte del repositorio de la Universidad de California, Irvine.

Su tarea consiste en crear un "Pandas dataframe" que contenga 559 filas (una por base de

datos) y las siguientes columnas:

- Nombre de la base de datos
- Link a la base de datos
- Tipo de datos
- Tipo de tarea a resolver (default task)
- Tipo de las variables
- Número de observaciones
- Número de variables
- Año
- Descripción de la base (Pista: Utilice la opción list view: https://archive.ics.uci.edu/ml/datasets.php?format=&task=&att=&area=&numAtt=&numIns=&type=&sort=nameUp&view=list)

In [116]:
```
import requests
from bs4 import BeautifulSoup
html = requests.get('https://archive.ics.uci.edu/ml/datasets.php').text
page = BeautifulSoup(html)
```

In [153]:
```
items = page.find_all('p')
```

In [154]:
```
import re
data= re.findall('class="normal".+', str(items))
```

In [155]:
```
nombres = re.findall('''>([\w'’  "0-9/?^.+\,;*$:!@()_+—&-<>⁄]*)</a></b
></p></td></tr></table></td>''', str(page))
nombres
```

Out[155]:
```
['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 &amp;amp; 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 (PU
C-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 co
llection',
 'NYSK',
 'Turkiye Student Evaluation',
 "ser Knowledge Modeling Data (Students' Knowledge Levels on DC Electri
cal Machines)",
 'EEG Eye State',
 'Physicochemical Properties of Protein Tertiary Structure',
 'seismic-bumps',
 'banknote authentication',
 'USPTO Algorithm Challenge, run by NASA-Harvard Tournament Lab and Top
Coder    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 Transit
ions',
 '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 P
arkinson’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 As
sisted 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 wris
tband 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 sp
eech',
 'Data for Software Engineering Teamwork Assessment in Education Settin
g',
 '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 magne
```

```
tometer',
 '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 we
arable sensor',
 'Immunotherapy Dataset',
 'Cryotherapy Dataset ',
 'OCT data &amp;amp; Color Fundus Images of Left &amp;amp; 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 sp
```

```
eech 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 cond
ition ',
 '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 cente
r',
 '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 var
iables',
```

```
 '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']
```

In [120]:
```
len(nombres)
```

Out[120]:
```
559
```

In [156]:
```
other= re.findall('class="normal">([A-Z \s a-z,-]*)', str(data))
other= other[9:]
other
```

Out[156]:
```
['Multivariate',
 'Classification',
 'Categorical, Integer, Real',
 '',
 '',
 '',
```

```
    '',
    'Multivariate',
    'Classification',
    'Categorical, Integer',
    '',
    '',
    '',
    '',
    '',
    'Multivariate',
    'Classification',
    'Categorical, Integer, Real',
    '',
    '',
    '',
    '',
    '',
    '',
    'Recommender-Systems',
    'Categorical',
    '',
    '',
    '',
    '',
    '',
    'Multivariate',
    'Classification',
    'Categorical, Integer, Real',
    '',
    '',
    '',
    '',
    '',
    'Multivariate',
    'Classification',
    'Categorical, Integer, Real',
    '',
    '',
    '',
    '',
    '',
    'Multivariate',
    'Classification',
    'Categorical',
```

```
                '',
                '',
                '',
                '',
                'Multivariate',
                'Classification',
                'Categorical',
                '',
                '',
                '',
                '',
                '',
                'Multivariate',
                'Regression',
                'Categorical, Real',
                '',
                '',
                '',
                '',
                '',
                'Multivariate',
                'Regression',
                'Categorical, Integer, Real',
                '',
                '',
                '',
                '',
                '',
                'Univariate, Text',
                'Classification',
                '',
                '',
                '',
                '',
                '',
                'Multivariate',
                'Classification',
                'Categorical',
                '',
                '',
                '',
                '',
                '',
```

```
                'Multivariate',
                'Classification',
                'Categorical',
                '',
                '',
                '',
                '',
                '',
                'Multivariate',
                'Classification',
                'Categorical',
                '',
                '',
                '',
                '',
                '',
                'Multivariate',
                'Classification',
                'Integer',
                '',
                '',
                '',
                '',
                '',
                'Multivariate',
                'Classification, Regression',
                'Real',
                '',
                '',
                '',
                '',
                '',
                'Multivariate',
                'Classification',
                'Real',
                '',
                '',
                '',
                '',
                '',
                'Multivariate',
                'Classification',
                'Categorical, Integer',
                '',
```

```
                '',
                '',
                '',
                'Multivariate',
                'Classification',
                'Categorical',
                '',
                '',
                '',
                '',
                '',
                'Multivariate',
                'Classification',
                'Categorical, Integer',
                '',
                '',
                '',
                '',
                '',
                'Multivariate, Data-Generator',
                'Classification',
                'Categorical, Integer',
                '',
                '',
                '',
                '',
                '',
                'Multivariate',
                'Classification',
                'Categorical',
                '',
                '',
                '',
                '',
                '',
                'Multivariate',
                'Classification',
                'Categorical, Integer',
                '',
                '',
                '',
                '',
                '',
                'Domain-Theory',
```

```
                '',
                '',
                '',
                '',
                '',
                '',
                '',
                'Univariate, Time-Series',
                '',
                'Categorical, Integer',
                '',
                '',
                '',
                '',
                '',
                'Multivariate, Spatial',
                'Classification',
                'Categorical',
                '',
                '',
                '',
                '',
                '',
                'Multivariate',
                'Classification',
                'Categorical, Integer, Real',
                '',
                '',
                '',
                '',
                '',
                'Multivariate, Domain-Theory',
                'Classification',
                'Categorical, Real, Integer',
                '',
                '',
                '',
                '',
                '',
                'Multivariate',
                'Regression',
                'Integer',
                '',
                '',
```

```
                '',
                '',
                'Multivariate',
                'Classification',
                'Categorical, Integer',
                '',
                '',
                '',
                '',
                'Multivariate',
                'Classification',
                'Categorical, Integer',
                '',
                '',
                '',
                '',
                'Multivariate',
                'Classification',
                'Categorical, Integer, Real',
                '',
                '',
                '',
                '',
                'Multivariate',
                'Classification',
                'Categorical, Integer',
                '',
                '',
                '',
                '',
                'Multivariate, Time-Series',
                '',
                'Categorical, Integer',
                '',
                '',
                '',
                '',
                'Data-Generator',
                '',
```

```
                    'Real',
                    '',
                    '',
                    '',
                    '',
                    '',
                    '',
                    '',
                    '',
                    '',
                    '',
                    '',
                    '',
                    '',
                    '',
                    '',
                    '',
                    '',
                    '',
                    'Multivariate',
                    'Classification',
                    'Categorical, Integer, Real',
                    '',
                    '',
                    '',
                    '',
                    '',
                    'Multivariate',
                    'Classification',
                    'Real',
                    '',
                    '',
                    '',
                    '',
                    'Multivariate',
                    'Classification',
                    'Categorical, Integer',
                    '',
                    '',
                    '',
```

```
            '',
            '',
            'Function-Learning',
            'Real',
            '',
            '',
            '',
            '',
            '',
            'Multivariate',
            'Classification',
            'Real',
            '',
            '',
            '',
            '',
            '',
            'Multivariate',
            'Classification',
            'Integer',
            '',
            '',
            '',
            '',
            '',
            'Multivariate',
            'Classification',
            'Categorical',
            '',
            '',
            '',
            '',
            '',
            'Multivariate',
            'Classification',
            'Categorical, Integer, Real',
            '',
            '',
            '',
            '',
            '',
            'Multivariate',
            'Classification',
            'Categorical, Integer, Real',
```

```
                '',
                '',
                '',
                '',
                '',
                'Multivariate',
                'Classification',
                'Categorical, Integer, Real',
                '',
                '',
                '',
                '',
                '',
                'Multivariate, Time-Series',
                '',
                'Real',
                '',
                '',
                '',
                '',
                'Multivariate',
                'Classification',
                'Real',
                '',
                '',
                '',
                '',
                '',
                'Multivariate',
                'Classification',
                'Categorical, Integer, Real',
                '',
                '',
                '',
                '',
                '',
                'Multivariate',
                'Classification',
                'Integer, Real',
                '',
                '',
                '',
                '',
                '',
```

```
                'Multivariate',
                'Classification',
                'Real',
                '',
                '',
                '',
                '',
                '',
                'Multivariate',
                'Classification',
                'Real',
                '',
                '',
                '',
                '',
                '',
                'Relational',
                'Relational-Learning',
                'Categorical',
                '',
                '',
                '',
                '',
                '',
                'Multivariate',
                '',
                'Categorical, Integer, Real',
                '',
                '',
                '',
                '',
                '',
                'Multivariate, Data-Generator',
                'Classification',
                'Categorical',
                '',
                '',
                '',
                '',
                '',
                'Multivariate',
                'Classification',
                'Categorical',
                '',
```

```
                '',
                '',
                '',
                'Multivariate',
                'Classification',
                'Integer',
                '',
                '',
                '',
                '',
                '',
                'Multivariate',
                '',
                'Categorical, Integer, Real',
                '',
                '',
                '',
                '',
                '',
                'Domain-Theory',
                '',
                '',
                '',
                '',
                '',
                '',
                '',
                'Multivariate',
                'Classification',
                'Integer',
                '',
                '',
                '',
                '',
                '',
                'Multivariate',
                'Classification',
                'Categorical',
                '',
                '',
                '',
                '',
                '',
                'Multivariate',
```

```
                'Classification',
                'Categorical, Integer, Real',
                '',
                '',
                '',
                '',
                '',
                'Multivariate',
                'Classification',
                'Categorical, Integer, Real',
                '',
                '',
                '',
                '',
                '',
                'Domain-Theory',
                '',
                'Categorical, Integer, Real',
                '',
                '',
                '',
                '',
                '',
                'Sequential, Domain-Theory',
                'Classification',
                'Categorical',
                '',
                '',
                '',
                '',
                '',
                'Sequential',
                'Classification',
                'Categorical',
                '',
                '',
                '',
                '',
                '',
                'Sequential, Domain-Theory',
                'Classification',
                'Categorical',
                '',
                '',
```

```
        '',
        '',
        'Multivariate',
        'Classification',
        'Categorical',
        '',
        '',
        '',
        '',
        'Domain-Theory',
        '',
        '',
        '',
        '',
        '',
        '',
        '',
        'Multivariate',
        'Classification',
        'Integer, Real',
        '',
        '',
        '',
        '',
        'Multivariate',
        'Classification',
        'Categorical',
        '',
        '',
        '',
        '',
        'Multivariate',
        'Classification',
        'Integer',
        '',
        '',
        '',
        '',
        'Multivariate',
        'Classification',
```

```
                'Integer',
                '',
                '',
                '',
                '',
                '',
                'Multivariate',
                'Classification',
                'Categorical',
                '',
                '',
                '',
                '',
                '',
                'Domain-Theory',
                '',
                '',
                '',
                '',
                '',
                '',
                '',
                'Multivariate',
                'Classification',
                'Integer, Real',
                '',
                '',
                '',
                '',
                '',
                'Multivariate',
                'Classification',
                'Integer',
                '',
                '',
                '',
                '',
                '',
                'Multivariate',
                'Classification',
                'Integer',
                '',
                '',
                '',
```

```
                '',
                'Multivariate',
                'Classification',
                'Categorical, Integer',
                '',
                '',
                '',
                '',
                '',
                '',
                'Multivariate',
                'Classification',
                'Categorical',
                '',
                '',
                '',
                '',
                '',
                'Domain-Theory',
                '',
                '',
                '',
                '',
                '',
                '',
                '',
                '',
                'Domain-Theory',
                '',
                '',
                '',
                '',
                '',
                '',
                '',
                '',
                'Multivariate, Data-Generator',
                'Classification',
                'Real',
                '',
                '',
                '',
                '',
                '',
                'Multivariate',
                'Regression',
                'Categorical, Integer',
```

```
                '',
                '',
                '',
                '',
                '',
                'Multivariate',
                'Classification',
                'Categorical',
                '',
                '',
                '',
                '',
                '',
                'Multivariate',
                'Regression',
                'Categorical',
                '',
                '',
                '',
                '',
                '',
                'Multivariate',
                'Classification',
                'Categorical',
                '',
                '',
                '',
                '',
                '',
                'Multivariate',
                'Classification',
                'Categorical',
                '',
                '',
                '',
                '',
                '',
                'Multivariate',
                'Regression',
                'Integer',
                '',
                '',
                '',
                '',
                '',
```

```
                'Multivariate',
                'Classification',
                'Integer, Real',
                '',
                '',
                '',
                '',
                'Multivariate',
                'Classification',
                'Integer, Real',
                '',
                '',
                '',
                '',
                'Multivariate',
                'Classification',
                'Categorical',
                '',
                '',
                '',
                '',
                'Multivariate',
                'Classification',
                'Integer',
                '',
                '',
                '',
                '',
                'Multivariate',
                'Clustering',
                'Categorical, Integer',
                '',
                '',
                '',
                '',
                '',
                '',
                '',
                '',
                '',
```

```
                '',
                '',
                '',
                '',
                'Domain-Theory',
                '',
                '',
                '',
                '',
                '',
                '',
                '',
                'Multivariate',
                'Classification',
                'Categorical, Integer',
                '',
                '',
                '',
                '',
                '',
                'Multivariate',
                'Classification',
                'Categorical',
                '',
                '',
                '',
                '',
                '',
                'Multivariate, Domain-Theory',
                'Classification',
                'Categorical, Real',
                '',
                '',
                '',
                '',
                '',
                'Multivariate',
                'Classification',
                'Categorical',
                '',
                '',
                '',
                '',
                '',
                'Multivariate',
```

```
                'Classification',
                'Categorical, Integer',
                '',
                '',
                '',
                '',
                '',
                'Multivariate',
                'Classification',
                'Categorical',
                '',
                '',
                '',
                '',
                '',
                'Multivariate',
                'Clustering',
                'Integer, Real',
                '',
                '',
                '',
                '',
                '',
                'Multivariate, Data-Generator',
                'Classification',
                'Real',
                '',
                '',
                '',
                '',
                '',
                'Multivariate, Data-Generator',
                'Classification',
                'Real',
                '',
                '',
                '',
                '',
                '',
                'Multivariate',
                'Classification',
                'Integer, Real',
                '',
                '',
```

```
                '',
                '',
                'Multivariate',
                'Classification',
                'Real',
                '',
                '',
                '',
                '',
                'Multivariate',
                'Classification',
                'Categorical, Integer',
                '',
                '',
                '',
                '',
                '',
                '',
                '',
                '',
                '',
                '',
                '',
                '',
                'Text',
                '',
                '',
                '',
                '',
                '',
                '',
                'Multivariate, Time-Series',
                'Classification',
                'Categorical, Real',
                '',
                '',
                '',
                '',
                'Multivariate, Time-Series',
                'Classification',
```

```
                'Real',
                '',
                '',
                '',
                '',
                '',
                'Multivariate',
                'Clustering',
                'Categorical',
                '',
                '',
                '',
                '',
                '',
                'Multivariate',
                'Classification',
                'Categorical, Integer',
                '',
                '',
                '',
                '',
                '',
                'Multivariate',
                '',
                'Categorical, Real',
                '',
                '',
                '',
                '',
                '',
                'Multivariate',
                '',
                'Real',
                '',
                '',
                '',
                '',
                '',
                'Relational',
                '',
                '',
                '',
                '',
                '',
                '',
```

```
'',
'Multivariate, Time-Series',
'',
'Categorical, Integer, Real',
'',
'',
'',
'',
'',
'Spatio-temporal',
'',
'Integer, Real',
'',
'',
'',
'',
'',
'Transactional, Sequential',
'Recommender-Systems',
'Categorical',
'',
'',
'',
'',
'',
'Image',
'Classification',
'Integer',
'',
'',
'',
'',
'',
'Multivariate',
'Regression, Description',
'Categorical, Integer',
'',
'',
'',
'',
'',
'Multivariate',
'',
'Categorical, Integer',
```

```
                '',
                '',
                '',
                '',
                '',
                'Multivariate',
                '',
                'Categorical, Integer',
                '',
                '',
                '',
                '',
                '',
                'Multivariate, Time-Series',
                'Classification',
                'Real',
                '',
                '',
                '',
                '',
                '',
                'Multivariate',
                'Regression',
                'Categorical, Integer',
                '',
                '',
                '',
                '',
                '',
                'Multivariate',
                'Classification',
                'Categorical, Integer',
                '',
                '',
                '',
                '',
                '',
                'Relational',
                '',
                '',
                '',
                '',
                '',
                '',
                '',
```

```
                'Multivariate, Relational',
                '',
                '',
                '',
                '',
                '',
                '',
                'Sequential',
                '',
                'Categorical',
                '',
                '',
                '',
                '',
                'Text',
                '',
                '',
                '',
                '',
                '',
                '',
                'Multivariate, Time-Series',
                '',
                'Categorical, Real',
                '',
                '',
                '',
                '',
                'Univariate, Time-Series',
                '',
                '',
                '',
                '',
                '',
                '',
                'Text',
                'Classification',
                'Categorical',
                '',
```

```
                '',
                '',
                '',
                '',
                'Multivariate, Time-Series',
                'Classification',
                'Integer',
                '',
                '',
                '',
                '',
                '',
                'Time-Series',
                'Classification, Clustering',
                'Real',
                '',
                '',
                '',
                '',
                '',
                'Multivariate, Text',
                'Classification',
                'Categorical',
                '',
                '',
                '',
                '',
                '',
                'Text, Sequential',
                '',
                '',
                '',
                '',
                '',
                '',
                '',
                'Image',
                'Classification',
                '',
                '',
                '',
                '',
                '',
                '',
                'Multivariate',
```

```
 'Classification',
 'Categorical, Integer, Real',
 '',
 '',
 '',
 '',
 '',
 'Multivariate',
 'Classification',
 'Categorical, Integer',
 '',
 '',
 '',
 '',
 '',
 'Multivariate',
 'Classification',
 'Categorical, Real',
 '',
 '',
 '',
 '',
 ...]
```

In [157]:
```
numbers = re.findall('class="normal">([0-9]*)', str(data))
numbers = numbers[12:]
numbers
```

Out[157]:
```
['4177',
 '8',
 '1995',
 '',
 '',
 '',
 '',
 '',
 '48842',
 '14',
 '1996',
 '',
 '',
 '',
 '',
 '',
```

```
                '798',
                '38',
                '',
                '',
                '',
                '',
                '',
                '',
                '37711',
                '294',
                '1998',
                '',
                '',
                '',
                '',
                '',
                '452',
                '279',
                '1998',
                '',
                '',
                '',
                '',
                '',
                '6000',
                '7',
                '1992',
                '',
                '',
                '',
                '',
                '',
                '226',
                '',
                '1987',
                '',
                '',
                '',
                '',
                '',
                '226',
                '69',
                '1992',
                '',
```

```
                '',
                '',
                '',
                '',
                '398',
                '8',
                '1993',
                '',
                '',
                '',
                '',
                '',
                '205',
                '26',
                '1987',
                '',
                '',
                '',
                '',
                '',
                '294',
                '1',
                '1994',
                '',
                '',
                '',
                '',
                '',
                '625',
                '4',
                '1994',
                '',
                '',
                '',
                '',
                '',
                '16',
                '4',
                '',
                '',
                '',
                '',
                '',
                '',
                '286',
```

```
                '9',
                '1988',
                '',
                '',
                '',
                '',
                '',
                '699',
                '10',
                '1992',
                '',
                '',
                '',
                '',
                '',
                '198',
                '34',
                '1995',
                '',
                '',
                '',
                '',
                '',
                '569',
                '32',
                '1995',
                '',
                '',
                '',
                '',
                '',
                '108',
                '13',
                '1990',
                '',
                '',
                '',
                '',
                '',
                '1728',
                '6',
                '1997',
                '',
                '',
```

```
                '',
                '',
                '48842',
                '14',
                '1996',
                '',
                '',
                '',
                '',
                '',
                '',
                '22',
                '1988',
                '',
                '',
                '',
                '',
                '',
                '3196',
                '36',
                '1989',
                '',
                '',
                '',
                '',
                '',
                '28056',
                '6',
                '1994',
                '',
                '',
                '',
                '',
                '',
                '',
                '',
                '',
                '',
                '',
                '',
                '',
                '',
                '',
                '',
                '100',
                '6',
```

```
                '',
                '',
                '',
                '',
                '',
                '',
                '67557',
                '42',
                '1995',
                '',
                '',
                '',
                '',
                '',
                '690',
                '15',
                '',
                '',
                '',
                '',
                '',
                '',
                '125',
                '',
                '1992',
                '',
                '',
                '',
                '',
                '',
                '209',
                '9',
                '1987',
                '',
                '',
                '',
                '',
                '',
                '1473',
                '9',
                '1997',
                '',
                '',
                '',
```

```
                '',
                '581012',
                '54',
                '1998',
                '',
                '',
                '',
                '',
                '',
                '512',
                '39',
                '1995',
                '',
                '',
                '',
                '',
                '',
                '366',
                '33',
                '1998',
                '',
                '',
                '',
                '',
                '',
                '',
                '20',
                '',
                '',
                '',
                '',
                '',
                '',
                '',
                '',
                '',
                '',
                '',
                '',
                '',
                '',
                '',
                '',
                '',
                '1994',
```

```
                '',
                '',
                '',
                '',
                '',
                '',
                '',
                '',
                '',
                '',
                '',
                '',
                '132',
                '12',
                '1989',
                '',
                '',
                '',
                '',
                '',
                '336',
                '8',
                '1996',
                '',
                '',
                '',
                '',
                '',
                '194',
                '30',
                '1990',
                '',
                '',
                '',
                '',
                '',
                '352',
                '',
                '1990',
                '',
                '',
                '',
                '',
                '',
```

```
                '214',
                '10',
                '1987',
                '',
                '',
                '',
                '',
                '',
                '306',
                '3',
                '1999',
                '',
                '',
                '',
                '',
                '',
                '160',
                '5',
                '1989',
                '',
                '',
                '',
                '',
                '',
                '303',
                '75',
                '1988',
                '',
                '',
                '',
                '',
                '',
                '155',
                '19',
                '1988',
                '',
                '',
                '',
                '',
                '',
                '368',
                '27',
                '1989',
                '',
```

```
                '',
                '',
                '',
                '',
                '',
                '',
                '',
                '',
                '',
                '',
                '',
                '',
                '2310',
                '19',
                '1990',
                '',
                '',
                '',
                '',
                '',
                '3279',
                '1558',
                '1998',
                '',
                '',
                '',
                '',
                '',
                '351',
                '34',
                '1989',
                '',
                '',
                '',
                '',
                '',
                '150',
                '4',
                '1988',
                '',
                '',
                '',
                '',
                '',
                '7797',
```

```
                '617',
                '1994',
                '',
                '',
                '',
                '',
                '',
                '104',
                '12',
                '1990',
                '',
                '',
                '',
                '',
                '',
                '57',
                '16',
                '1988',
                '',
                '',
                '',
                '',
                '',
                '',
                '7',
                '1988',
                '',
                '',
                '',
                '',
                '',
                '24',
                '4',
                '1990',
                '',
                '',
                '',
                '',
                '',
                '20000',
                '16',
                '1991',
                '',
                '',
```

```
            '',
            '',
            '345',
            '7',
            '1990',
            '',
            '',
            '',
            '',
            '',
            '',
            '',
            '',
            '',
            '',
            '',
            '',
            '32',
            '56',
            '1992',
            '',
            '',
            '',
            '',
            '148',
            '18',
            '1988',
            '',
            '',
            '',
            '',
            '209',
            '8',
            '1990',
            '',
            '',
            '',
            '',
            '528',
            '22',
```

```
                '1996',
                '',
                '',
                '',
                '',
                '',
                '',
                '',
                '1995',
                '',
                '',
                '',
                '',
                '',
                '106',
                '58',
                '1990',
                '',
                '',
                '',
                '',
                '',
                '128',
                '',
                '',
                '',
                '',
                '',
                '',
                '',
                '3190',
                '61',
                '1992',
                '',
                '',
                '',
                '',
                '',
                '432',
                '7',
                '1992',
                '',
                '',
                '',
```

```
                '',
                '202',
                '',
                '1994',
                '',
                '',
                '',
                '',
                '',
                '2000',
                '649',
                '',
                '',
                '',
                '',
                '',
                '',
                '8124',
                '22',
                '1987',
                '',
                '',
                '',
                '',
                '',
                '476',
                '168',
                '1994',
                '',
                '',
                '',
                '',
                '',
                '6598',
                '168',
                '1994',
                '',
                '',
                '',
                '',
                '',
                '12960',
                '8',
                '1997',
```

```
                '',
                '',
                '',
                '',
                '',
                '',
                '',
                '1991',
                '',
                '',
                '',
                '',
                '',
                '5473',
                '10',
                '1995',
                '',
                '',
                '',
                '',
                '',
                '5620',
                '64',
                '1998',
                '',
                '',
                '',
                '',
                '',
                '10992',
                '16',
                '1998',
                '',
                '',
                '',
                '',
                '',
                '90',
                '8',
                '1993',
                '',
                '',
                '',
                '',
                '',
```

```
                '339',
                '17',
                '1988',
                '',
                '',
                '',
                '',
                '',
                '',
                '',
                '',
                '',
                '',
                '',
                '',
                '',
                '',
                '',
                '',
                '',
                '',
                '',
                '',
                '',
                '',
                '',
                '',
                '',
                '',
                '',
                '72',
                '1992',
                '',
                '',
                '',
                '',
                '',
                '167',
                '4',
                '1993',
                '',
                '',
                '',
                '',
                '',
                '15',
                '6',
                '1988',
                '',
```

```
                '',
                '',
                '',
                '',
                '1389',
                '10',
                '1989',
                '',
                '',
                '',
                '',
                '',
                '307',
                '35',
                '1988',
                '',
                '',
                '',
                '',
                '',
                '47',
                '35',
                '1987',
                '',
                '',
                '',
                '',
                '',
                '23',
                '4',
                '1993',
                '',
                '',
                '',
                '',
                '',
                '531',
                '102',
                '1988',
                '',
                '',
                '',
                '',
                '',
                '4601',
```

```
                '57',
                '1999',
                '',
                '',
                '',
                '',
                '',
                '267',
                '22',
                '2001',
                '',
                '',
                '',
                '',
                '',
                '267',
                '44',
                '2001',
                '',
                '',
                '',
                '',
                '',
                '76',
                '45',
                '',
                '',
                '',
                '',
                '',
                '',
                '',
                '',
                '1992',
                '',
                '',
                '',
                '',
                '',
                '1000',
                '',
                '1993',
                '',
                '',
```

```
                '',
                '',
                '151',
                '5',
                '1997',
                '',
                '',
                '',
                '',
                '',
                '958',
                '9',
                '1991',
                '',
                '',
                '',
                '',
                '',
                '7200',
                '21',
                '1987',
                '',
                '',
                '',
                '',
                '',
                '10',
                '32',
                '1994',
                '',
                '',
                '',
                '',
                '',
                '285',
                '17',
                '1988',
                '',
                '',
                '',
                '',
                '',
                '435',
                '16',
```

```
                '1987',
                '',
                '',
                '',
                '',
                '',
                '527',
                '38',
                '1993',
                '',
                '',
                '',
                '',
                '',
                '5000',
                '21',
                '1988',
                '',
                '',
                '',
                '',
                '',
                '5000',
                '40',
                '1988',
                '',
                '',
                '',
                '',
                '',
                '178',
                '13',
                '1991',
                '',
                '',
                '',
                '',
                '',
                '1484',
                '8',
                '1996',
                '',
                '',
                '',
```

```
                '',
                '101',
                '17',
                '1990',
                '',
                '',
                '',
                '',
                '',
                '',
                '',
                '',
                '',
                '',
                '',
                '',
                '',
                '20000',
                '',
                '1999',
                '',
                '',
                '',
                '',
                '',
                '6650',
                '15',
                '1999',
                '',
                '',
                '',
                '',
                '',
                '2565',
                '22',
                '2002',
                '',
                '',
                '',
                '',
                '',
                '2458285',
                '68',
                '',
```

```
                '',
                '',
                '',
                '',
                '',
                '299285',
                '40',
                '2000',
                '',
                '',
                '',
                '',
                '',
                '',
                '340',
                '17',
                '1999',
                '',
                '',
                '',
                '',
                '',
                '',
                '68040',
                '89',
                '1999',
                '',
                '',
                '',
                '',
                '',
                '',
                '',
                '',
                '2001',
                '',
                '',
                '',
                '',
                '',
                '',
                '122',
                '4',
                '1999',
                '',
                '',
                '',
                '',
                '',
```

```
                '178080',
                '12',
                '1999',
                '',
                '',
                '',
                '',
                '',
                '50672',
                '',
                '2000',
                '',
                '',
                '',
                '',
                '',
                '640',
                '',
                '1999',
                '',
                '',
                '',
                '',
                '',
                '9000',
                '86',
                '2000',
                '',
                '',
                '',
                '',
                '',
                '10104',
                '72',
                '1999',
                '',
                '',
                '',
                '',
                '',
                '256932',
                '61',
                '1999',
                '',
```

```
                '',
                '',
                '',
                '',
                '640',
                '12',
                '',
                '',
                '',
                '',
                '',
                '',
                '191779',
                '481',
                '1998',
                '',
                '',
                '',
                '',
                '',
                '4000000',
                '42',
                '1999',
                '',
                '',
                '',
                '',
                '',
                '',
                '',
                '',
                '2001',
                '',
                '',
                '',
                '',
                '',
                '10000',
                '',
                '1999',
                '',
                '',
                '',
                '',
                '',
                '989818',
```

```
                '',
                '',
                '',
                '',
                '',
                '',
                '',
                '129000',
                '',
                '2003',
                '',
                '',
                '',
                '',
                '',
                '',
                '',
                '1999',
                '',
                '',
                '',
                '',
                '',
                '100000',
                '',
                '1999',
                '',
                '',
                '',
                '',
                '',
                '21578',
                '5',
                '1997',
                '',
                '',
                '',
                '',
                '',
                '463',
                '90',
                '1999',
                '',
                '',
```

```
                '',
                '',
                '600',
                '',
                '1999',
                '',
                '',
                '',
                '',
                '332',
                '5',
                '1998',
                '',
                '',
                '',
                '',
                '',
                '',
                '',
                '',
                '',
                '',
                '',
                '',
                '',
                '',
                '',
                '',
                '',
                '',
                '690',
                '14',
                '',
                '',
                '',
                '',
                '',
                '',
                '1000',
                '20',
```

```
 '1994',
 '',
 '',
 '',
 '',
 '270',
 '13',
 '',
 '',
 '',
 '',
 '',
 ...]
```

In [158]:
```
data_types =[]
iterable = []
for i in range(3914):
    if i % 7 == 0:
        iterable.append(i)
for i in iterable:
    data_types.append(other[i])
data_types=data_types[:559]
data_types
```

Out[158]:
```
['Multivariate',
 'Multivariate',
 'Multivariate',
 '',
 'Multivariate',
 'Multivariate',
 'Multivariate',
 'Multivariate',
 'Multivariate',
 'Multivariate',
 'Univariate, Text',
 'Multivariate',
 'Multivariate',
 'Multivariate',
 'Multivariate',
 'Multivariate',
 'Multivariate'
```

```
'Multivariate',
'Multivariate',
'Multivariate',
'Multivariate',
'Multivariate, Data-Generator',
'Multivariate',
'Multivariate',
'Domain-Theory',
'Univariate, Time-Series',
'Multivariate, Spatial',
'Multivariate',
'Multivariate, Domain-Theory',
'Multivariate',
'Multivariate',
'Multivariate',
'Multivariate',
'Multivariate',
'Multivariate, Time-Series',
'Data-Generator',
'',
'',
'Multivariate',
'Multivariate',
'Multivariate',
'',
'Multivariate',
'Multivariate',
'Multivariate',
'Multivariate',
'Multivariate',
'Multivariate',
'Multivariate, Time-Series',
'Multivariate',
'Multivariate',
'Multivariate',
'Multivariate',
'Multivariate',
'Multivariate',
'Relational',
'Multivariate',

'Multivariate, Data-Generator'
```

```
'Multivariate, Data-Generator',
'Multivariate',
'Multivariate',
'Multivariate',
'Domain-Theory',
'Multivariate',
'Multivariate',
'Multivariate',
'Multivariate',
'Domain-Theory',
'Sequential, Domain-Theory',
'Sequential',
'Sequential, Domain-Theory',
'Multivariate',
'Domain-Theory',
'Multivariate',
'Multivariate',
'Multivariate',
'Multivariate',
'Multivariate',
'Domain-Theory',
'Multivariate',
'Multivariate',
'Multivariate',
'Multivariate',
'Multivariate',
'Domain-Theory',
'Domain-Theory',
'Multivariate, Data-Generator',
'Multivariate',
'Multivariate',
'Multivariate',
'Multivariate',
'Multivariate',
'Multivariate',
'Multivariate',
'Multivariate',
'Multivariate',
'Multivariate',
'Multivariate',

'Multivariate'
```

```
'Multivariate',
'',
'Domain-Theory',
'Multivariate',
'Multivariate',
'Multivariate, Domain-Theory',
'Multivariate',
'Multivariate',
'Multivariate',
'Multivariate',
'Multivariate, Data-Generator',
'Multivariate, Data-Generator',
'Multivariate',
'Multivariate',
'Multivariate',
'',
'Text',
'Multivariate, Time-Series',
'Multivariate, Time-Series',
'Multivariate',
'Multivariate',
'Multivariate',
'Multivariate',
'Relational',
'Multivariate, Time-Series',
'Spatio-temporal',
'Transactional, Sequential',
'Image',
'Multivariate',
'Multivariate',
'Multivariate',
'Multivariate, Time-Series',
'Multivariate',
'Multivariate',
'Relational',
'Multivariate, Relational',
'Sequential',
'Text',
'Multivariate, Time-Series',

'Univariate, Time-Series'
```

```
'Univariate, Time-Series',
'Text',
'Multivariate, Time-Series',
'Time-Series',
'Multivariate, Text',
'Text, Sequential',
'Image',
'Multivariate',
'Multivariate',
'Multivariate',
'Multivariate',
'Multivariate',
'Multivariate',
'Multivariate',
'Multivariate',
'Multivariate',
'',
'Domain-Theory',
'',
'Multivariate',
'Multivariate, Time-Series',
'Multivariate, Time-Series',
'Multivariate',
'Multivariate',
'Multivariate, Sequential',
'Multivariate',
'Multivariate',
'Text',
'Text',
'Multivariate',
'Sequential',
'Multivariate',
'Multivariate',
'Multivariate',
'Multivariate',
'Multivariate',
'Multivariate, Sequential, Time-Series',
'Multivariate',
'Multivariate',

'Time-Series'
```

```
 'Time-Series',
'Multivariate',
'Multivariate, Sequential',
'Multivariate',
'Multivariate',
'Multivariate',
'Multivariate, Sequential',
'Multivariate',
'Multivariate',
'Multivariate',
'Multivariate',
'Multivariate, Time-Series',
'Multivariate',
'Multivariate',
'Multivariate',
'Text',
'Multivariate',
'Multivariate',
'Multivariate, Sequential',
'Multivariate, Time-Series',
'Univariate, Sequential, Time-Series',
'Multivariate',
'Multivariate',
'Multivariate',
'Multivariate',
'Multivariate, Time-Series',
'Text',
'Domain-Theory',
'Multivariate, Sequential',
'Multivariate',
'Multivariate',
'Multivariate',
'Multivariate',
'Time-Series',
'Time-Series',
'Multivariate, Text, Domain-Theory',
'Time-Series, Domain-Theory',
'Multivariate, Text, Domain-Theory',
'Text',

'Text'
```

```
 'Text',
'Multivariate, Univariate, Text',
'Multivariate, Univariate, Text',
'Multivariate',
'Text',
'Multivariate',
'Multivariate',
'Multivariate, Time-Series',
'Univariate',
'Multivariate, Text, Domain-Theory',
'Univariate',
'Univariate',
'Multivariate, Time-Series',
'Multivariate',
'Multivariate, Text',
'Multivariate, Time-Series',
'Multivariate',
'Multivariate, Univariate, Text',
'Sequential',
'Text',
'Multivariate, Time-Series',
'',
'Multivariate',
'Multivariate',
'Multivariate',
'Multivariate, Time-Series',
'Sequential, Text',
'Multivariate, Univariate, Time-Series',
'Time-Series, Multivariate',
'Multivariate',
'Sequential',
'Multivariate, Time-Series',
'Multivariate',
'Multivariate',
'Multivariate',
'Multivariate',
'Multivariate, Time-Series',
'Multivariate',
'Multivariate',

'Multivariate, Sequential, Text'
```

```
'Multivariate, Sequential, Text',
'Multivariate',
'Multivariate',
'Multivariate, Sequential, Time-Series',
'Multivariate',
'Multivariate',
'Multivariate',
'Domain-Theory',
'Multivariate, Text',
'Multivariate, Time-Series',
'Multivariate, Sequential, Time-Series',
'Multivariate',
'Multivariate',
'Multivariate, Sequential, Time-Series, Text',
'Univariate',
'Multivariate, Sequential, Time-Series',
'Multivariate',
'Multivariate, Time-Series',
'',
'',
'Multivariate',
'Multivariate',
'Multivariate, Time-Series',
'Multivariate',
'Univariate, Sequential, Time-Series',
'Univariate, Sequential, Time-Series',
'Multivariate',
'Text',
'Multivariate',
'Multivariate',
'Multivariate',
'Text',
'Multivariate',
'Multivariate',
'Multivariate',
'Sequential',
'Multivariate',
'Multivariate',
'Multivariate',

'Multivariate, Sequential, Time-Series'
```

```
'Multivariate, Sequential, Time-Series',
'Univariate, Domain-Theory',
'Multivariate',
'Multivariate, Time-Series',
'',
'Multivariate',
'Multivariate, Time-Series',
'Multivariate, Time-Series',
'Multivariate',
'Text',
'Time-Series',
'Time-Series',
'Multivariate',
'Multivariate',
'Multivariate',
'Multivariate, Sequential',
'Multivariate',
'Multivariate, Time-Series',
'Multivariate',
'Time-Series',
'Multivariate, Time-Series',
'Multivariate, Univariate, Sequential, Text',
'Multivariate',
'Multivariate',
'Multivariate',
'',
'Multivariate, Time-Series',
'Multivariate',
'Multivariate',
'Text',
'Multivariate',
'Multivariate',
'Multivariate',
'Multivariate',
'Multivariate',
'Sequential, Time-Series',
'Multivariate',
'Multivariate, Sequential, Time-Series, Domain-Theory',
'Multivariate',

'Multivariate, Time-Series'
```

```
'Multivariate, Time-Series',
'Multivariate',
'Multivariate, Sequential, Time-Series',
'Multivariate, Time-Series',
'Multivariate, Sequential, Time-Series',
'Multivariate',
'Multivariate, Sequential, Time-Series',
'Multivariate, Sequential, Time-Series',
'Multivariate',
'Multivariate',
'Multivariate, Sequential, Time-Series',
'Multivariate',
'Multivariate',
'Multivariate',
'Multivariate, Time-Series',
'Multivariate',
'Multivariate',
'Multivariate, Time-Series',
'Multivariate, Time-Series, Domain-Theory',
'Multivariate, Time-Series',
'Multivariate',
'Time-Series',
'Multivariate',
'Multivariate, Sequential, Time-Series',
'Multivariate',
'Multivariate',
'Multivariate',
'Text',
'Multivariate',
'Multivariate',
'Multivariate, Time-Series',
'Text',
'Multivariate, Text',
'Multivariate, Sequential, Time-Series',
'Multivariate',
'Multivariate',
'Text',
'Multivariate, Time-Series',
'Multivariate, Sequential',

'Multivariate'
```

```
'Multivariate',
'Multivariate',
'Multivariate, Time-Series',
'Multivariate, Time-Series',
'Multivariate, Time-Series',
'',
'Multivariate',
'Multivariate',
'Multivariate',
'Sequential, Time-Series',
'Multivariate, Time-Series',
'Multivariate',
'Multivariate, Time-Series',
'',
'Text',
'Multivariate',
'Multivariate',
'Multivariate',
'Multivariate, Sequential, Time-Series',
'',
'Text',
'Multivariate',
'Multivariate',
'Text',
'Multivariate',
'Time-Series',
'Text',
'',
'',
'Multivariate, Time-Series',
'Multivariate',
'Sequential, Text',
'Multivariate, Text',
'Multivariate',
'Text',
'Multivariate',
'Multivariate',
'Multivariate',

'Multivariate'
```

```
'Multivariate',
'Multivariate',
'Multivariate',
'Text',
'',
'Sequential',
'Univariate',
'Univariate',
'Multivariate',
'Multivariate',
'Multivariate, Time-Series, Text',
'Multivariate',
'Multivariate',
'Multivariate, Sequential, Time-Series',
'Univariate, Domain-Theory',
'Multivariate',
'Text',
'Sequential',
'Multivariate',
'Multivariate',
'Multivariate, Sequential',
'Multivariate, Time-Series',
'Multivariate',
'Multivariate, Time-Series',
'Univariate',
'Multivariate',
'Multivariate, Text',
'Multivariate',
'Multivariate, Time-Series',
'Multivariate',
'Text',
'Time-Series',
'Multivariate, Text',
'Multivariate, Time-Series',
'Text',
'Multivariate',
'Multivariate',
'Multivariate, Text',
'Multivariate, Text',

'Multivariate'
```

```
'Multivariate',
'Multivariate, Time-Series',
'Multivariate, Sequential, Time-Series',
'Multivariate',
'Multivariate',
'Univariate',
'Multivariate',
'Multivariate',
'Univariate',
'Time-Series',
'Time-Series',
'Multivariate',
'Multivariate, Text',
'Multivariate',
'Multivariate',
'',
'Multivariate',
'Time-Series',
'Multivariate, Univariate, Sequential, Time-Series',
'Multivariate, Time-Series',
'Multivariate, Text',
'Multivariate, Text',
'Multivariate',
'Multivariate, Time-Series',
'Multivariate',
'Multivariate',
'Multivariate, Sequential, Time-Series',
'Multivariate',
'Multivariate',
'Multivariate, Time-Series',
'Multivariate',
'Multivariate, Univariate',
'Multivariate, Sequential',
'Text',
'Time-Series',
'Multivariate, Time-Series',
'Multivariate, Sequential, Time-Series, Text',
'Multivariate',

'Multivariate'
```

```
'Multivariate',
'Multivariate, Sequential, Time-Series',
'Multivariate, Time-Series',
'Multivariate',
'Multivariate',
'Multivariate',
'Multivariate',
'Multivariate, Text',
'Multivariate, Time-Series, Text',
'Multivariate',
'Multivariate, Time-Series',
'Multivariate, Sequential, Time-Series',
'Multivariate',
'Multivariate',
'Multivariate',
'Multivariate',
'Time-Series',
'Multivariate',
'Multivariate',
'Multivariate',
'Multivariate, Time-Series',
'Multivariate, Time-Series',
'Multivariate',
'Multivariate',
'Multivariate',
'Text',
'Multivariate',
'Multivariate',
'Multivariate, Time-Series, Text',
'Multivariate',
'Multivariate, Text',
'Multivariate, Sequential, Time-Series',
'Multivariate, Univariate',
'Multivariate',
'Multivariate, Univariate',
'Multivariate, Time-Series',
'Multivariate',
'Multivariate',

''
```

```
 ,
'Multivariate',
'Multivariate',
'Multivariate',
'Multivariate',
'Multivariate, Time-Series',
'Multivariate, Time-Series',
'Univariate',
'Multivariate',
'Multivariate',
'Multivariate, Sequential',
'Sequential, Time-Series',
'Multivariate',
'Multivariate',
'Multivariate',
'Multivariate',
'Multivariate, Time-Series',
'Multivariate',
'Multivariate',
'Multivariate',
'Multivariate',
'Multivariate',
'Multivariate',
'Multivariate',
'Multivariate',
'Multivariate',
'Multivariate',
'Text',
'Multivariate',
'Multivariate',
'Multivariate',
'',
'Text',
'Text',
'Multivariate',
'Time-Series']
```

In [124]:
```
len(data_types)
```

Out[124]: 559

In [125]:
```
default_task=[]
iterablee = []
for i in range(3914):
    if i % 7 == 0:
        iterablee.append(i+1)
iterablee.remove(3914)
for i in iterablee:
    default_task.append(other[i])
default_task
```

Out[125]:
```
['Classification',
 'Classification',
 'Classification',
 'Recommender-Systems',
 'Classification',
 'Classification',
 'Classification',
 'Classification',
 'Regression',
 'Regression',
 'Classification',
 'Classification',
 'Classification',
 'Classification',
 'Classification',
 'Classification, Regression',
 'Classification',
 'Classification',
 'Classification',
 'Classification',
 'Classification',
 'Classification',
 'Classification',
 '',
 '',
 'Classification',
 'Classification',
 'Classification',
```

```
                'Regression',
                'Classification',
                'Classification',
                'Classification',
                'Classification',
                '',
                '',
                '',
                '',
                'Classification',
                'Classification',
                'Classification',
                'Function-Learning',
                'Classification',
                'Classification',
                'Classification',
                'Classification',
                'Classification',
                'Classification',
                '',
                'Classification',
                'Classification',
                'Classification',
                'Classification',
                'Classification',
                'Relational-Learning',
                '',
                'Classification',
                'Classification',
                'Classification',
                '',
                '',
                'Classification',
                'Classification',
                'Classification',
                'Classification',
                '',
                'Classification',
                'Classification',
```

```
                'Classification',
                'Classification',
                '',
                'Classification',
                'Classification',
                'Classification',
                'Classification',
                'Classification',
                '',
                'Classification',
                'Classification',
                'Classification',
                'Classification',
                'Classification',
                '',
                '',
                'Classification',
                'Regression',
                'Classification',
                'Regression',
                'Classification',
                'Classification',
                'Regression',
                'Classification',
                'Classification',
                'Classification',
                'Classification',
                'Clustering',
                '',
                '',
                'Classification',
                'Classification',
                'Classification',
                'Classification',
                'Classification',
                'Classification',
                'Clustering',
                'Classification',
                'Classification',
```

```
                'Classification',
                'Classification',
                'Classification',
                '',
                '',
                'Classification',
                'Classification',
                'Clustering',
                'Classification',
                '',
                '',
                '',
                '',
                '',
                '',
                'Recommender-Systems',
                'Classification',
                'Regression, Description',
                '',
                '',
                'Classification',
                'Regression',
                'Classification',
                '',
                '',
                '',
                '',
                '',
                '',
                '',
                'Classification',
                'Classification',
                'Classification, Clustering',
                'Classification',
                '',
                'Classification',
                'Classification',
                'Classification',
                'Classification',
                'Classification',
                'Classification',
```

```
                'Classification',
                'Classification',
                '',
                'Classification',
                'Classification',
                '',
                '',
                '',
                '',
                '',
                'Classification',
                'Classification',
                'Classification',
                'Classification',
                'Regression',
                'Classification',
                'Clustering',
                'Regression',
                'Classification',
                'Classification',
                'Classification',
                'Classification',
                'Classification',
                'Classification',
                'Classification',
                'Causal-Discovery',
                'Classification',
                'Classification, Clustering',
                'Classification',
                'Classification',
                'Classification',
                'Classification, Causal-Discovery',
                'Clustering',
                'Classification, Clustering',
                'Regression',
                'Regression',
                'Classification',
                'Classification, Regression',
                'Classification',
```

```
'Classification',
'Regression',
'Classification',
'',
'Classification',
'Classification',
'Classification',
'Classification',
'Classification',
'Classification',
'Classification',
'Classification',
'Regression',
'Classification',
'',
'Regression',
'Classification',
'Classification',
'Classification',
'Regression',
'Classification',
'Classification',
'Classification',
'Classification',
'Regression, Clustering, Causal-Discovery',
'Classification, Clustering',
'Classification',
'Classification',
'Classification, Regression, Clustering',
'Classification, Regression, Clustering',
'Classification',
'Classification',
'Classification',
'Classification',
'Classification',
'Classification',
'Classification, Clustering',
'Classification',
'Classification',
```

```
                'Classification',
                '',
                'Classification',
                'Regression, Clustering',
                'Classification, Clustering',
                'Classification',
                '',
                'Classification',
                'Classification, Clustering',
                'Classification',
                'Classification, Regression',
                'Regression',
                'Classification, Regression',
                'Classification',
                'Regression, Clustering',
                'Classification, Regression',
                'Regression, Classification',
                'Classification',
                'Classification',
                'Classification',
                'Classification',
                'Classification',
                'Classification',
                'Classification',
                'Classification, Clustering',
                'Classification, Clustering',
                'Classification',
                'Clustering',
                'Classification, Clustering',
                'Classification',
                'Classification',
                'Regression',
                'Classification',
                'Classification',
                'Classification',
                'Classification, Clustering',
                'Classification, Regression, Clustering, Causa',
                'Classification, Clustering',
                'Regression',
```

```
                'Classification',
                'Regression',
                'Regression',
                '',
                'Classification',
                '',
                'Classification',
                'Classification',
                'Classification',
                'Classification',
                'Classification, Clustering',
                'Classification',
                'Classification, Clustering',
                'Classification, Clustering',
                'Classification',
                'Classification, Clustering',
                'Classification, Regression, Clustering',
                'Regression',
                'Classification, Clustering',
                'Classification',
                'Regression',
                'Classification',
                'Classification, Clustering',
                'Classification',
                'Classification, Clustering, Causal-Discovery',
                'Classification, Regression, Clustering',
                'Classification, Regression',
                'Classification, Clustering',
                'Classification, Clustering',
                'Regression',
                'Classification',
                'Classification',
                'Clustering',
                'Classification, Regression',
                'Classification, Regression',
                'Classification, Regression',
                'Classification',
                'Classification, Clustering',
                'Classification',
```

```
                'Clustering',
                'Classification, Regression',
                'Regression',
                'Classification, Clustering',
                'Classification, Regression',
                'Classification',
                'Classification, Regression',
                'Regression, Clustering',
                'Classification, Regression',
                'Classification, Causal-Discovery',
                'Classification',
                'Classification',
                'Classification, Clustering',
                'Classification',
                'Regression',
                'Classification',
                'Classification',
                'Classification',
                'Classification, Regression',
                'Classification',
                'Regression, Clustering, Causal-Discovery',
                'Regression',
                'Classification',
                'Classification',
                'Classification, Clustering',
                'Clustering, Causal-Discovery',
                'Classification, Regression',
                'Classification',
                'Classification, Clustering',
                'Classification, Regression, Clustering',
                'Classification, Clustering',
                'Classification, Regression, Clustering',
                'Classification',
                'Classification',
                'Classification, Regression, Clustering',
                'Classification',
                'Classification',
                'Classification, Clustering',
                'Causal-Discovery',
```

```
'Classification, Regression',
'Classification',
'Classification',
'Classification, Regression, Clustering',
'Classification, Clustering',
'Regression',
'Classification, Regression',
'Classification',
'Regression',
'Classification',
'Classification',
'Classification',
'Classification',
'Regression',
'Causal-Discovery',
'Clustering',
'Classification, Clustering',
'Classification',
'Regression',
'Classification, Clustering, Causal-Discovery',
'Classification, Regression',
'Classification, Regression, Clustering',
'Classification, Regression, Clustering',
'Classification',
'Classification',
'Regression',
'Classification, Regression',
'Classification',
'Classification',
'Causal-Discovery',
'Classification, Clustering',
'Regression',
'Classification, Clustering',
'Classification',
'Regression',
'Classification, Clustering',
'Classification, Regression',
'Classification',
'Regression',
```

```
                'Classification, Regression, Clustering',
                'Clustering',
                'Classification, Regression',
                '',
                'Classification',
                'Classification',
                'Classification, Clustering',
                'Classification',
                'Classification, Clustering',
                'Classification',
                'Classification, Clustering',
                'Classification, Clustering',
                'Classification, Clustering',
                'Classification',
                'Regression',
                'Classification, Regression',
                'Classification',
                'Classification',
                'Classification, Regression',
                'Classification',
                'Clustering',
                'Clustering',
                'Classification',
                'Classification',
                'Classification',
                'Classification',
                'Classification',
                'Classification',
                'Classification',
                'Classification, Regression',
                'Classification',
                'Classification',
                'Classification',
                'Classification',
                'Classification',
                'Classification',
                'Classification',
                'Regression',
                'Classification',
```

```
'Classification, Clustering',
'Classification, Clustering',
'Classification, Regression',
'Regression',
'Clustering',
'Classification',
'Regression',
'Clustering',
'Classification, Clustering',
'Classification, Clustering',
'Classification, Clustering',
'Classification, Regression',
'Regression',
'Classification, Regression',
'Classification',
'Classification',
'Regression',
'Classification, Clustering',
'Classification',
'Classification',
'Classification',
'Classification, Regression',
'Classification',
'Classification',
'Recommendation',
'Classification, Regression, Clustering',
'Classification, Regression, Clustering',
'Classification',
'Regression',
'Classification, Regression',
'Regression',
'Classification',
'Classification, Clustering',
'Classification',
'Classification',
'Classification, Regression',
'Classification',
'Classification',
'Classification',
```

```
'Classification',
'Classification, Clustering',
'Regression',
'Classification',
'Classification',
'Classification',
'Classification',
'Classification, Regression, Clustering',
'Classification, Regression',
'Classification, Clustering',
'Classification, Clustering',
'Classification',
'Classification, Regression',
'Clustering',
'Classification',
'Regression',
'Regression, Clustering',
'Regression',
'Regression',
'Classification, Regression, Clustering',
'Classification',
'Regression, Clustering',
'Classification',
'Classification, Clustering',
'Regression',
'Classification, Regression, Clustering',
'Classification',
'Regression',
'Regression',
'Classification',
'Classification',
'Classification',
'Classification',
'Classification, Regression',
'Regression',
'Classification, Clustering',
'Regression',
'Regression',
'Classification, Regression',
```

```
'Classification, Clustering, Causal-Discovery',
'Classification',
'Classification',
'Classification, Regression, Clustering',
'Classification',
'Classification, Clustering',
'Classification, Regression, Clustering',
'Classification, Clustering',
'Classification',
'Classification',
'Classification, Clustering',
'Classification',
'Classification',
'Classification',
'Classification',
'Clustering',
'Regression',
'Classification',
'Regression',
'Classification',
'Classification, Regression, Causal-Discovery',
'Classification, Clustering',
'Classification',
'Classification',
'Classification, Clustering',
'Classification',
'Classification',
'Classification',
'Classification, Regression, Clustering',
'Classification',
'Clustering',
'Classification, Regression',
'Classification, Clustering',
'Causal-Discovery',
'Clustering',
'Regression, Clustering',
'Classification',
'Classification, Regression, Clustering',
'Classification, Regression',
```

```
 'Classification, Regression, Clustering',
 'Classification, Clustering',
 'Classification',
 'Classification',
 'Classification, Regression',
 'Regression',
 'Classification',
 'Classification, Clustering',
 'Classification, Regression',
 'Classification',
 'Classification, Regression',
 'Classification, Clustering',
 'Classification',
 'Classification',
 'Classification',
 'Classification',
 'Classification, Clustering',
 'Classification',
 'Classification, Regression, Clustering',
 'Regression',
 'Classification',
 'Classification',
 'Classification, Clustering',
 'Classification']
```

In [126]:
```
len(default_task)
```

Out[126]: 559

In [127]:
```
attribute_types=[]
iterableee = []
for i in range(3914):
    if i % 7 == 0:
        iterableee.append(i+2)
iterableee.remove(3915)
for i in iterableee:
    attribute_types.append(other[i])
attribute_types
```

```
Out[127]: ['Categorical, Integer, Real',
           'Categorical, Integer',
           'Categorical, Integer, Real',
           'Categorical',
           'Categorical, Integer, Real',
           'Categorical, Integer, Real',
           'Categorical',
           'Categorical',
           'Categorical, Real',
           'Categorical, Integer, Real',
           '',
           'Categorical',
           'Categorical',
           'Categorical',
           'Integer',
           'Real',
           'Real',
           'Categorical, Integer',
           'Categorical',
           'Categorical, Integer',
           'Categorical, Integer',
           'Categorical',
           'Categorical, Integer',
           '',
           'Categorical, Integer',
           'Categorical',
           'Categorical, Integer, Real',
           'Categorical, Real, Integer',
           'Integer',
           'Categorical, Integer',
           'Categorical, Integer',
           'Categorical, Integer, Real',
           'Categorical, Integer',
           'Categorical, Integer',
           'Real',
           '',
           '',
           'Categorical, Integer, Real',
           'Real',
```

```
                'Categorical, Integer',
                'Real',
                'Real',
                'Integer',
                'Categorical',
                'Categorical, Integer, Real',
                'Categorical, Integer, Real',
                'Categorical, Integer, Real',
                'Real',
                'Real',
                'Categorical, Integer, Real',
                'Integer, Real',
                'Real',
                'Real',
                'Categorical',
                'Categorical, Integer, Real',
                'Categorical',
                'Categorical',
                'Integer',
                'Categorical, Integer, Real',
                '',
                'Integer',
                'Categorical',
                'Categorical, Integer, Real',
                'Categorical, Integer, Real',
                'Categorical, Integer, Real',
                'Categorical',
                'Categorical',
                'Categorical',
                'Categorical',
                '',
                'Integer, Real',
                'Categorical',
                'Integer',
                'Integer',
                'Categorical',
                '',
                'Integer, Real',
                'Integer',
```

```
'Integer',
'Categorical, Integer',
'Categorical',
'',
'',
'Real',
'Categorical, Integer',
'Categorical',
'Categorical',
'Categorical',
'Categorical',
'Integer',
'Integer, Real',
'Integer, Real',
'Categorical',
'Integer',
'Categorical, Integer',
'',
'',
'Categorical, Integer',
'Categorical',
'Categorical, Real',
'Categorical',
'Categorical, Integer',
'Categorical',
'Integer, Real',
'Real',
'Real',
'Integer, Real',
'Real',
'Categorical, Integer',
'',
'',
'Categorical, Real',
'Real',
'Categorical',
'Categorical, Integer',
'Categorical, Real',
'Real',
```

```
'',
'Categorical, Integer, Real',
'Integer, Real',
'Categorical',
'Integer',
'Categorical, Integer',
'Categorical, Integer',
'Categorical, Integer',
'Real',
'Categorical, Integer',
'Categorical, Integer',
'',
'',
'Categorical',
'',
'Categorical, Real',
'',
'Categorical',
'Integer',
'Real',
'Categorical',
'',
'',
'Categorical, Integer, Real',
'Categorical, Integer',
'Categorical, Real',
'Integer',
'Real',
'Integer',
'Integer',
'Categorical',
'Real',
'Real',
'',
'',
'Real',
'Categorical, Integer',
'Categorical, Integer',
'Categorical, Integer',
```

```
                'Real',
                'Integer',
                'Integer',
                'Real',
                '',
                'Integer',
                'Real',
                'Real',
                'Real',
                'Integer',
                'Integer',
                'Integer',
                'Real',
                'Real',
                'Integer',
                'Real',
                'Real',
                'Real',
                'Integer',
                'Integer',
                'Real',
                'Categorical',
                'Real',
                'Real',
                'Real',
                'Categorical, Integer',
                'Real',
                'Integer, Real',
                'Real',
                'Integer, Real',
                'Integer',
                '',
                'Real',
                'Real',
                'Real',
                'Real',
                'Real',
                'Categorical, Integer, Real',
                'Integer, Real',
```

```
                'Real',
                'Real',
                'Real',
                '',
                'Real',
                'Integer',
                'Integer, Real',
                'Real',
                'Real',
                'Real',
                'Real',
                'Real',
                'Real',
                '',
                'Real',
                '',
                '',
                'Integer, Real',
                'Integer, Real',
                'Real',
                '',
                'Real',
                'Integer, Real',
                'Real',
                'Real',
                'Real',
                'Real',
                'Real',
                'Real',
                '',
                'Integer',
                'Real',
                'Real',
                'Integer, Real',
                'Integer',
                '',
                '',
                'Real',
                'Integer, Real',
```

```
                'Real',
                'Real',
                'Real',
                'Real',
                'Real',
                'Integer, Real',
                'Real',
                'Integer, Real',
                'Real',
                'Real',
                'Real',
                'Integer, Real',
                '',
                'Real',
                'Integer',
                'Real',
                '',
                '',
                'Real',
                'Integer, Real',
                'Real',
                'Real',
                'Real',
                'Integer',
                'Integer, Real',
                'Real',
                '',
                'Integer, Real',
                'Real',
                'Real',
                'Integer, Real',
                'Integer, Real',
                'Integer, Real',
                'Real',
                'Real',
                'Real',
                '',
                'Real',
                '',
```

```
            '',
            'Real',
            'Real',
            'Real',
            '',
            'Real',
            'Real',
            'Integer',
            '',
            'Real',
            '',
            'Integer',
            '',
            'Real',
            'Integer, Real',
            'Integer, Real',
            'Real',
            'Integer',
            'Integer, Real',
            'Real',
            '',
            '',
            'Real',
            'Real',
            'Integer, Real',
            'Integer',
            'Integer, Real',
            'Real',
            '',
            'Real',
            'Real',
            'Real',
            'Real',
            'Real',
            'Integer',
            'Real',
            'Real',
            'Integer, Real',
            '',
```

```
                'Real',
                'Real',
                'Integer',
                'Real',
                'Integer, Real',
                'Categorical',
                '',
                'Integer, Real',
                '',
                '',
                'Integer, Real',
                'Real',
                'Real',
                '',
                'Real',
                'Real',
                'Real',
                'Real',
                'Integer, Real',
                'Real',
                'Integer',
                'Real',
                'Real',
                'Integer',
                'Integer, Real',
                'Real',
                'Integer, Real',
                'Integer',
                'Real',
                'Real',
                'Real',
                'Real',
                '',
                'Real',
                'Real',
                'Real',
                'Integer, Real',
                'Real',
                'Real',
```

```
                'Real',
                '',
                'Integer',
                '',
                'Integer',
                'Real',
                'Real',
                'Real',
                'Integer',
                'Integer',
                'Integer, Real',
                'Real',
                'Integer',
                '',
                'Integer, Real',
                'Integer',
                'Integer, Real',
                'Real',
                'Integer',
                'Real',
                'Real',
                'Integer, Real',
                'Integer',
                'Real',
                'Integer, Real',
                'Integer, Real',
                'Integer, Real',
                'Integer, Real',
                'Integer',
                'Integer, Real',
                'Integer',
                '',
                'Integer, Real',
                '',
                'Real',
                'Real',
                '',
                'Integer',
                'Real',
```

```
                'Real',
                'Integer',
                'Real',
                'Integer',
                'Integer',
                'Integer, Real',
                'Integer, Real',
                'Integer',
                'Integer',
                'Real',
                'Integer',
                'Integer, Real',
                'Integer',
                'Integer',
                'Integer',
                'Integer, Real',
                'Real',
                'Integer, Real',
                'Integer',
                '',
                'Integer',
                'Real',
                'Integer, Real',
                'Integer, Real',
                'Real',
                '',
                'Integer, Real',
                'Real',
                '',
                'Integer',
                'Real',
                'Real',
                'Real',
                'Integer',
                'Integer',
                'Real',
                'Real',
                'Integer, Real',
                '',
```

```
                'Real',
                'Real',
                'Integer, Real',
                'Integer',
                'Integer',
                'Real',
                '',
                '',
                'Integer',
                'Integer',
                'Integer',
                '',
                'Real',
                'Categorical',
                'Integer',
                'Integer',
                'Integer',
                'Real',
                'Real',
                'Real',
                '',
                'Integer, Real',
                '',
                'Integer, Real',
                'Real',
                'Integer',
                '',
                '',
                'Real',
                'Real',
                'Integer, Real',
                'Real',
                'Integer',
                'Real',
                'Real',
                'Real',
                'Integer, Real',
                'Real',
                'Real',
```

```
                'Integer',
                'Real',
                'Integer',
                '',
                'Integer, Real',
                'Real',
                'Real',
                'Real',
                'Real',
                'Integer',
                'Integer',
                'Integer',
                'Real',
                'Integer, Real',
                'Integer, Real',
                'Integer, Real',
                'Real',
                'Real',
                'Integer, Real',
                'Real',
                '',
                '',
                '',
                'Integer, Real',
                'Integer',
                'Integer, Real',
                'Real',
                'Real',
                'Real',
                'Real',
                'Real',
                'Integer, Real',
                'Real',
                'Real',
                'Integer, Real',
                'Integer',
                'Real',
                'Real',
                'Integer, Real',
```

```
                '',
                'Integer, Real',
                '',
                '',
                'Integer, Real',
                'Integer',
                '',
                'Real',
                '',
                'Real',
                'Integer',
                '',
                'Integer',
                'Integer, Real',
                'Integer, Real',
                '',
                '',
                'Integer',
                'Real',
                'Integer, Real',
                'Real',
                'Real',
                'Integer',
                'Real',
                'Real',
                'Real',
                'Integer, Real',
                'Real',
                '',
                '',
                '',
                'Real',
                'Real',
                'Integer, Real',
                'Real',
                '',
                'Integer',
                'Real',
                'Integer, Real',
```

```
 'Integer',
 '',
 'Integer',
 'Real',
 '',
 'Integer, Real',
 'Integer',
 'Integer, Real',
 'Integer',
 '',
 '',
 '',
 'Real']
```

In [128]:
```
len(attribute_types)
```

Out[128]:
```
559
```

In [129]:
```
num_instances =[]
iiterable = []
for i in range(3911):
    if i % 7 == 0:
        iiterable.append(i)
for i in iiterable:
    num_instances.append(numbers[i])
num_instances
```

Out[129]:
```
['4177',
 '48842',
 '798',
 '37711',
 '452',
 '6000',
 '226',
 '226',
 '398',
 '205',
 '294',
 '625',
```

```
'16',
'286',
'699',
'198',
'569',
'108',
'1728',
'48842',
'',
'3196',
'28056',
'',
'100',
'67557',
'690',
'125',
'209',
'1473',
'581012',
'512',
'366',
'',
'',
'',
'',
'132',
'336',
'194',
'352',
'214',
'306',
'160',
'303',
'155',
'368',
'',
'2310',
'3279',
'351',
```

```
                '150',
                '7797',
                '104',
                '57',
                '',
                '24',
                '20000',
                '345',
                '',
                '32',
                '148',
                '209',
                '528',
                '',
                '106',
                '128',
                '3190',
                '432',
                '202',
                '2000',
                '8124',
                '476',
                '6598',
                '12960',
                '',
                '5473',
                '5620',
                '10992',
                '90',
                '339',
                '',
                '',
                '',
                '167',
                '15',
                '1389',
                '307',
                '47',
                '23',
```

```
'531',
'4601',
'267',
'267',
'76',
'',
'1000',
'151',
'958',
'7200',
'10',
'285',
'435',
'527',
'5000',
'5000',
'178',
'1484',
'101',
'',
'20000',
'6650',
'2565',
'2458285',
'299285',
'340',
'68040',
'',
'122',
'178080',
'50672',
'640',
'9000',
'10104',
'256932',
'640',
'191779',
'4000000',
'',
```

```
                '10000',
                '989818',
                '129000',
                '',
                '100000',
                '21578',
                '463',
                '600',
                '332',
                '',
                '',
                '690',
                '1000',
                '270',
                '6435',
                '2310',
                '58000',
                '946',
                '20008',
                '208',
                '528',
                '',
                '',
                '1024',
                '10080',
                '50400',
                '1025010',
                '19020',
                '1364',
                '961',
                '517',
                '200',
                '8000000',
                '1030',
                '606',
                '900',
                '2600',
                '1950',
                '13500',
```

```
                '4400',
                '2536',
                '300',
                '197',
                '2858',
                '748',
                '11640',
                '1593',
                '1567',
                '22632',
                '360',
                '103',
                '1994',
                '120',
                '4898',
                '2396130',
                '16772',
                '5875',
                '503',
                '51',
                '106',
                '2126',
                '5456',
                '8800',
                '164860',
                '',
                '1941',
                '130065',
                '515345',
                '440',
                '',
                '53500',
                '8235',
                '',
                '5749132',
                '2215',
                '310',
                '10000',
                '3000',
```

```
                '1500',
                '30000',
                '2500',
                '4143',
                '64',
                '53414',
                '65554',
                '45211',
                '1138562',
                '13910',
                '583',
                '2551',
                '34465',
                '5574',
                '245057',
                '182',
                '3850505',
                '138',
                '1080',
                '2075259',
                '210',
                '115',
                '3960456',
                '',
                '10299',
                '1600',
                '768',
                '308',
                '100',
                '237',
                '434874',
                '536',
                '140000',
                '6118',
                '165632',
                '18000',
                '540',
                '931',
                '1055',
```

```
                '100',
                '9120',
                '403',
                '111740',
                '10421',
                '5820',
                '403',
                '14980',
                '45730',
                '2584',
                '1372',
                '306',
                '120000',
                '13910',
                '2747',
                '3395',
                '39242',
                '4137',
                '17389',
                '51',
                '470',
                '132',
                '5000000',
                '11000000',
                '250',
                '126',
                '',
                '4889',
                '',
                '',
                '340',
                '501',
                '45781',
                '1503',
                '440',
                '2000',
                '9568',
                '168',
                '100000',
```

```
'5665',
'79',
'127',
'1040',
'9900',
'560',
'60021',
'1419',
'101',
'399',
'58',
'180',
'21048',
'',
'750',
'3000',
'150',
'1059',
'11934',
'27965',
'216',
'120',
'649',
'370',
'4178504',
'221579',
'10800',
'58509',
'129685',
'2456',
'2921',
'1151',
'6590',
'3000',
'39797',
'326',
'913',
'168286',
'400',
```

```
                '314080',
                '637',
                '1710671',
                '12000',
                '10929',
                '1080',
                '40000',
                '43930257',
                '230318',
                '10500000',
                '13197',
                '',
                '30000',
                '324',
                '541909',
                '11164866',
                '163',
                '373',
                '20560',
                '40',
                '422937',
                '9358',
                '640',
                '919438',
                '40949',
                '5744',
                '10503',
                '42240',
                '102944',
                '500',
                '9782222',
                '11463',
                '17898',
                '1885',
                '19735',
                '1540',
                '4007',
                '153540',
                '606',
```

```
                '1353',
                '1956',
                '43824',
                '3942',
                '858',
                '287',
                '17764280',
                '106574',
                '9358',
                '11500',
                '92000',
                '315',
                '78095',
                '130',
                '74',
                '52854',
                '77',
                '811',
                '504',
                '401',
                '2856',
                '10546',
                '801',
                '1540',
                '1451',
                '1075',
                '78095',
                '7195',
                '3600',
                '76',
                '60',
                '405',
                '303',
                '303',
                '107888',
                '76000',
                '10000',
                '180',
                '745000',
```

```
                '12234',
                '292',
                '104',
                '60000',
                '2000',
                '165',
                '217',
                '1175',
                '704',
                '75128',
                '90',
                '90',
                '50',
                '71',
                '93239',
                '540',
                '105',
                '6611',
                '15',
                '372',
                '58000',
                '4960',
                '241600',
                '130000',
                '7062606',
                '740',
                '70',
                '2205',
                '10721',
                '640',
                '1000',
                '116',
                '1672',
                '322',
                '93600',
                '3060',
                '5879',
                '9200',
                '20000',
```

```
                '20867',
                '',
                '4143',
                '215063',
                '6000000',
                '21263',
                '63000000',
                '10190',
                '300',
                '12330',
                '5180',
                '756',
                '10000',
                '80',
                '1184',
                '1047',
                '777',
                '249',
                '414',
                '',
                '143',
                '7840',
                '30000',
                '35717',
                '135',
                '980',
                '5456',
                '120',
                '4095000',
                '7051',
                '240',
                '48204',
                '260000',
                '288000',
                '8300000',
                '125',
                '170',
                '141712',
                '3916',
```

```
'6262',
'420768',
'1067371',
'1385',
'908',
'546',
'13956534',
'15630426',
'8992',
'1687',
'779',
'1056',
'590',
'21643',
'7750',
'14057567',
'27170754',
'597',
'329',
'299',
'20000',
'26136',
'1000',
'399',
'24017',
'325834',
'2916697',
'22470',
'189',
'520',
'826',
'2279',
'28764',
'7107',
'288000',
'9800',
'4760',
'72',
'1450',
```

```
                '170',
                '1984',
                '2955',
                '65532',
                '65919',
                '2111',
                '3810',
                '18',
                '244',
                '104',
                '139',
                '360177',
                '36733',
                '4480',
                '165474',
                '1985',
                '10000',
                '232',
                '150',
                '11',
                '14057567',
                '8760',
                '48',
                '6321',
                '3150',
                '17256',
                '187',
                '399',
                '14',
                '467',
                '597',
                '23700',
                '615',
                '6819',
                '1000',
                '',
                '2633',
                '200',
```

```
 '13028',
 '800']
```

In [130]:
```
len(num_instances)
```

Out[130]:
```
559
```

In [131]:
```
num_attributes=[]
iiterablee = []
for i in range(3911):
    if i % 7 == 0:
        iiterablee.append(i+1)

for i in iiterablee:
    num_attributes.append(numbers[i])
num_attributes
```

Out[131]:
```
['8',
 '14',
 '38',
 '294',
 '279',
 '7',
 '',
 '69',
 '8',
 '26',
 '1',
 '4',
 '4',
 '9',
 '10',
 '34',
 '32',
 '13',
 '6',
 '14',
 '22',
 '36',
```

```
                '6',
                '',
                '6',
                '42',
                '15',
                '',
                '9',
                '9',
                '54',
                '39',
                '33',
                '20',
                '',
                '',
                '',
                '12',
                '8',
                '30',
                '',
                '10',
                '3',
                '5',
                '75',
                '19',
                '27',
                '',
                '19',
                '1558',
                '34',
                '4',
                '617',
                '12',
                '16',
                '7',
                '4',
                '16',
                '7',
                '',

                '56',
```

```
                '18',
                '8',
                '22',
                '',
                '58',
                '',
                '61',
                '7',
                '',
                '649',
                '22',
                '168',
                '168',
                '8',
                '',
                '10',
                '64',
                '16',
                '8',
                '17',
                '',
                '',
                '72',
                '4',
                '6',
                '10',
                '35',
                '35',
                '4',
                '102',
                '57',
                '22',
                '44',
                '45',
                '',
                '',
                '5',
                '9',

                '21',
```

```
'32',
'17',
'16',
'38',
'21',
'40',
'13',
'8',
'17',
'',
'',
'15',
'22',
'68',
'40',
'17',
'89',
'',
'4',
'12',
'',
'',
'86',
'72',
'61',
'12',
'481',
'42',
'',
'',
'',
'',
'',
'',
'5',
'90',
'',
'5',

'',
```

```
'',
'14',
'20',
'13',
'36',
'19',
'9',
'18',
'4',
'60',
'10',
'',
'',
'10',
'4',
'3',
'11',
'11',
'',
'6',
'13',
'',
'100000',
'9',
'101',
'10000',
'20000',
'100000',
'5000',
'500',
'73',
'43',
'23',
'3',
'5',
'',
'256',
'591',

'70',
```

```
                '91',
                '10',
                '128',
                '6',
                '12',
                '3231961',
                '5409',
                '26',
                '',
                '',
                '10',
                '23',
                '24',
                '13',
                '8',
                '',
                '27',
                '50',
                '90',
                '138672',
                '',
                '386',
                '',
                '',
                '12',
                '147',
                '6',
                '8',
                '27',
                '10000',
                '20000',
                '10000',
                '54877',
                '4702',
                '24',
                '29',
                '17',
                '3',

                '128',
```

```
                '10',
                '242',
                '120',
                '',
                '4',
                '13',
                '52',
                '47',
                '857',
                '9',
                '7',
                '200',
                '4',
                '',
                '561',
                '64',
                '8',
                '7',
                '10',
                '9',
                '4',
                '8',
                '77',
                '51',
                '18',
                '1950000',
                '18',
                '1300',
                '41',
                '6',
                '5625',
                '5',
                '',
                '7',
                '33',
                '5',
                '15',
                '9',

                '19',
```

```
                '5',
                '5',
                '1000000',
                '129',
                '',
                '20',
                '152',
                '24',
                '16',
                '35',
                '17',
                '5',
                '18',
                '28',
                '7',
                '309',
                '3',
                '6',
                '',
                '',
                '16',
                '13',
                '5',
                '6',
                '8',
                '2',
                '4',
                '148',
                '55',
                '17',
                '8',
                '42',
                '26',
                '50',
                '2',
                '281',
                '120',
                '',

                '6',
```

```
                '120432',
                '150000',
                '529',
                '',
                '16',
                '2500',
                '5',
                '68',
                '16',
                '100',
                '216',
                '23',
                '33',
                '140256',
                '19',
                '20',
                '20',
                '49',
                '12',
                '30',
                '5232',
                '20',
                '1',
                '',
                '61',
                '27',
                '53',
                '11',
                '25',
                '0',
                '20',
                '9',
                '3',
                '561',
                '82',
                '13',
                '16',
                '13',

                '28',
```

```
                '4',
                '',
                '24',
                '34',
                '8',
                '128',
                '15',
                '513',
                '7',
                '7',
                '5',
                '15',
                '480000',
                '11',
                '54',
                '561',
                '64',
                '6',
                '116',
                '19',
                '',
                '5812',
                '9',
                '32',
                '29',
                '67',
                '',
                '25',
                '6400',
                '10',
                '5',
                '13',
                '98',
                '36',
                '69',
                '2158859',
                '518',
                '15',

                '179',
```

```
'',
'12',
'38',
'65',
'102',
'86',
'7',
'53',
'20',
'1',
'71',
'29',
'20531',
'65',
'3',
'22',
'38',
'22',
'4814',
'698',
'13',
'10',
'59',
'56',
'482',
'171',
'5',
'500',
'411',
'8519',
'21',
'21',
'171',
'7',
'49',
'12',
'3',
'21',

'9',
```

```
                '8',
                '7',
                '2',
                '11',
                '11',
                '173',
                '5',
                '15',
                '3',
                '105',
                '25000',
                '',
                '18',
                '21000',
                '115',
                '21',
                '206',
                '43680',
                '8',
                '10',
                '59',
                '10',
                '5',
                '5',
                '1000',
                '138',
                '',
                '16',
                '2',
                '10',
                '',
                '8',
                '6',
                '129',
                '81',
                '12',
                '6',
                '22',

                '18',
```

```
                '9',
                '754',
                '14',
                '5',
                '',
                '',
                '',
                '18',
                '7',
                '7',
                '',
                '',
                '7',
                '5',
                '6',
                '4',
                '18',
                '11',
                '25',
                '',
                '20',
                '12',
                '46',
                '9',
                '8',
                '49',
                '11',
                '8',
                '54',
                '36',
                '3916',
                '710',
                '18',
                '8',
                '29',
                '7',
                '9',
                '37',
                '6',
                '1024',

                '1024',
```

```
                '14',
                '7',
                '8265',
                '29',
                '25',
                '3',
                '115',
                '1',
                '12',
                '13',
                '200000',
                '6',
                '21',
                '4',
                '2400',
                '175',
                '10',
                '4714',
                '23',
                '17',
                '2',
                '9',
                '8',
                '280',
                '49',
                '3',
                '14',
                '19',
                '',
                '54',
                '8',
                '1087',
                '12',
                '3',
                '17',
                '8',
                '9',
                '12',

                '1656',
```

```
'6',
'2',
'11',
'533',
'14',
'84',
'22',
'16',
'52',
'19',
'3',
'14',
'321',
'13',
'13',
'55',
'39',
'4',
'7',
'79',
'1',
'',
'14',
'96',
'21',
'',
'5',
'2',
'69',
'9']
```

In [132]:
```
len(num_attributes)
```

Out[132]: 559

In [133]:
```
year=[]
iiterableee = []
for i in range(3911):
    if i % 7 == 0:
```

```
        iiterableee.append(i+2)
for i in iiterableee:
    year.append(numbers[i])
year
```

Out[133]:
```
['1995',
 '1996',
 '',
 '1998',
 '1998',
 '1992',
 '1987',
 '1992',
 '1993',
 '1987',
 '1994',
 '1994',
 '',
 '1988',
 '1992',
 '1995',
 '1995',
 '1990',
 '1997',
 '1996',
 '1988',
 '1989',
 '1994',
 '',
 '',
 '1995',
 '',
 '1992',
 '1987',
 '1997',
 '1998',
 '1995',
 '1998',
 '',
 '',
```

```
                '1994',
                '',
                '1989',
                '1996',
                '1990',
                '1990',
                '1987',
                '1999',
                '1989',
                '1988',
                '1988',
                '1989',
                '',
                '1990',
                '1998',
                '1989',
                '1988',
                '1994',
                '1990',
                '1988',
                '1988',
                '1990',
                '1991',
                '1990',
                '',
                '1992',
                '1988',
                '1990',
                '1996',
                '1995',
                '1990',
                '',
                '1992',
                '1992',
                '1994',
                '',
                '1987',
                '1994',
                '1994',
```

```
                '1997',
                '1991',
                '1995',
                '1998',
                '1998',
                '1993',
                '1988',
                '',
                '',
                '1992',
                '1993',
                '1988',
                '1989',
                '1988',
                '1987',
                '1993',
                '1988',
                '1999',
                '2001',
                '2001',
                '',
                '1992',
                '1993',
                '1997',
                '1991',
                '1987',
                '1994',
                '1988',
                '1987',
                '1993',
                '1988',
                '1988',
                '1991',
                '1996',
                '1990',
                '',
                '1999',
                '1999',
                '2002',
```

```
                '',
                '2000',
                '1999',
                '1999',
                '2001',
                '1999',
                '1999',
                '2000',
                '1999',
                '2000',
                '1999',
                '1999',
                '',
                '1998',
                '1999',
                '2001',
                '1999',
                '',
                '2003',
                '1999',
                '1999',
                '1997',
                '1999',
                '1999',
                '1998',
                '',
                '',
                '',
                '1994',
                '',
                '1993',
                '1990',
                '',
                '',
                '',
                '',
                '',
                '',
                '',
                '',
```

```
                '1989',
                '2006',
                '2006',
                '2007',
                '2007',
                '2007',
                '2007',
                '2008',
                '2008',
                '2008',
                '2007',
                '2008',
                '2008',
                '2008',
                '2008',
                '2008',
                '2008',
                '2008',
                '2008',
                '2008',
                '2008',
                '2009',
                '2008',
                '2008',
                '2008',
                '2009',
                '2009',
                '2009',
                '2009',
                '2009',
                '2009',
                '2010',
                '2009',
                '2010',
                '2010',
                '2010',
                '2010',
                '2010',
```

```
              '2010',
              '2010',
              '2010',
              '2010',
              '2010',
              '2011',
              '2011',
              '2011',
              '2011',
              '2011',
              '2011',
              '2011',
              '2011',
              '2011',
              '2011',
              '2011',
              '2011',
              '2011',
              '2011',
              '2011',
              '2011',
              '2011',
              '2011',
              '2012',
              '2012',
              '2012',
              '2012',
              '2012',
              '2012',
              '2012',
              '2012',
              '2012',
              '2012',
              '2012',
              '2012',
              '2012',
              '2012',
              '2012',
              '2012',
```

```
                 '2012',
                 '2012',
                 '2012',
                 '2012',
                 '2013',
                 '2013',
                 '2013',
                 '2013',
                 '2013',
                 '2013',
                 '2013',
                 '2013',
                 '2013',
                 '2013',
                 '2013',
                 '2013',
                 '2013',
                 '2013',
                 '2013',
                 '2013',
                 '2013',
                 '2013',
                 '2013',
                 '2013',
                 '2013',
                 '2013',
                 '2013',
                 '2013',
                 '2013',
                 '2013',
                 '2013',
                 '2013',
                 '2013',
                 '2014',
                 '2013',
                 '2013',
                 '2013',
                 '2014',
                 '2014',
```

```
'2014',
'2014',
'2014',
'2014',
'2014',
'2014',
'2014',
'2014',
'2014',
'2013',
'2014',
'2014',
'2014',
'2014',
'2014',
'2014',
'2014',
'2014',
'2014',
'2014',
'2014',
'2014',
'2014',
'2014',
'2014',
'2014',
'2014',
'2014',
'2014',
'2014',
'2014',
'2014',
'2014',
'2014',
'2014',
'2015',
'2014',
'2014',
```

```
                '2015',
                '2015',
                '2015',
                '2015',
                '2015',
                '2015',
                '2015',
                '2015',
                '2014',
                '2015',
                '2015',
                '2015',
                '2015',
                '2015',
                '2015',
                '2015',
                '2015',
                '2015',
                '2015',
                '2015',
                '2015',
                '2015',
                '2015',
                '2015',
                '2015',
                '2015',
                '2016',
                '2016',
                '2015',
                '2016',
                '2016',
                '2015',
                '2016',
                '2016',
                '2016',
                '2016',
                '2016',
                '2016',
                '2016',
                '2016',
```

```
                '2016',
                '2016',
                '2016',
                '2016',
                '2016',
                '2016',
                '2016',
                '2016',
                '2016',
                '2017',
                '2016',
                '2017',
                '2016',
                '2017',
                '2017',
                '2016',
                '2016',
                '2017',
                '2017',
                '2016',
                '2017',
                '2017',
                '2017',
                '2017',
                '2016',
                '2017',
                '2016',
                '2016',
                '2016',
                '2017',
                '2017',
                '2017',
                '2017',
                '2017',
                '2017',
                '2017',
                '2017',
                '2016',
                '2016',
                '2016',
```

```
                '2016',
                '2017',
                '2017',
                '2017',
                '2017',
                '2017',
                '2016',
                '2017',
                '2017',
                '2017',
                '2017',
                '2017',
                '2017',
                '2017',
                '2016',
                '2016',
                '2016',
                '2017',
                '2017',
                '2017',
                '2017',
                '2017',
                '2017',
                '2017',
                '2017',
                '2016',
                '2018',
                '2018',
                '2016',
                '2018',
                '2018',
                '2018',
                '2018',
                '2018',
                '2018',
                '2018',
                '2018',
                '2018',
                '2018',
```

```
                '2018',
                '2018',
                '2018',
                '2018',
                '2018',
                '2018',
                '2018',
                '2018',
                '2018',
                '2018',
                '2018',
                '2018',
                '2018',
                '2018',
                '2018',
                '2018',
                '2018',
                '2018',
                '2018',
                '2018',
                '2018',
                '2018',
                '2018',
                '2018',
                '2018',
                '2018',
                '2018',
                '2018',
                '2018',
                '2018',
                '2018',
                '2018',
                '2018',
                '2018',
                '2018',
                '2018',
                '2018',
                '2018',
                '2018',
                '2019',
```

```
                '2019',
                '2018',
                '2018',
                '2018',
                '2019',
                '2019',
                '2019',
                '2019',
                '2019',
                '2019',
                '2019',
                '2019',
                '2019',
                '2019',
                '2019',
                '2019',
                '2019',
                '2019',
                '2019',
                '2019',
                '2019',
                '2019',
                '2019',
                '2019',
                '2019',
                '2019',
                '2019',
                '2019',
                '2019',
                '2019',
                '2019',
                '2020',
                '2020',
                '2019',
                '2020',
                '2020',
                '2020',
                '2020',
                '2020',
                '2019',
```

```
                '2020',
                '2020',
                '2020',
                '2020',
                '2020',
                '2020',
                '2020',
                '2019',
                '2019',
                '2019',
                '2019',
                '2019',
                '2019',
                '2019',
                '2019',
                '2019',
                '2019',
                '2019',
                '2019',
                '2019',
                '2019',
                '2019',
                '2019',
                '2019',
                '2019',
                '2019',
                '2019',
                '2019',
                '2019',
                '2019',
                '2019',
                '2019',
                '2019',
                '2020',
                '2020',
                '2020',
                '2020',
                '2020',
                '2020',
```

```
 '2020',
 '2020',
 '2020',
 '2020',
 '2020',
 '2020',
 '2020',
 '2020',
 '2020',
 '2020',
 '2020',
 '2020',
 '2020',
 '2020',
 '2020',
 '2020',
 '2020']
```

In [134]:
```
len(year)
```

Out[134]:
```
559
```

In [135]:
```
links = re.findall("<td><p class=\"normal\"><b><a href=\"([/\%=\w'+():.
 ,-]*)\"", str(page))
link=[]
for i in links:
    link.append("https://archive.ics.uci.edu/ml/"+i)
link
```

Out[135]:
```
['https://archive.ics.uci.edu/ml/datasets/Abalone',
 'https://archive.ics.uci.edu/ml/datasets/Adult',
 'https://archive.ics.uci.edu/ml/datasets/Annealing',
 'https://archive.ics.uci.edu/ml/datasets/Anonymous+Microsoft+Web+Dat
a',
 'https://archive.ics.uci.edu/ml/datasets/Arrhythmia',
 'https://archive.ics.uci.edu/ml/datasets/Artificial+Characters',
 'https://archive.ics.uci.edu/ml/datasets/Audiology+%28Original%29',
 'https://archive.ics.uci.edu/ml/datasets/Audiology+%28Standardized%2
9',
```

```
 'https://archive.ics.uci.edu/ml/datasets/Auto+MPG',
 'https://archive.ics.uci.edu/ml/datasets/Automobile',
 'https://archive.ics.uci.edu/ml/datasets/Badges',
 'https://archive.ics.uci.edu/ml/datasets/Balance+Scale',
 'https://archive.ics.uci.edu/ml/datasets/Balloons',
 'https://archive.ics.uci.edu/ml/datasets/Breast+Cancer',
 'https://archive.ics.uci.edu/ml/datasets/Breast+Cancer+Wisconsin+%28Or
iginal%29',
 'https://archive.ics.uci.edu/ml/datasets/Breast+Cancer+Wisconsin+%28Pr
ognostic%29',
 'https://archive.ics.uci.edu/ml/datasets/Breast+Cancer+Wisconsin+%28Di
agnostic%29',
 'https://archive.ics.uci.edu/ml/datasets/Pittsburgh+Bridges',
 'https://archive.ics.uci.edu/ml/datasets/Car+Evaluation',
 'https://archive.ics.uci.edu/ml/datasets/Census+Income',
 'https://archive.ics.uci.edu/ml/datasets/Chess+%28King-Rook+vs.+King-K
night%29',
 'https://archive.ics.uci.edu/ml/datasets/Chess+%28King-Rook+vs.+King-P
awn%29',
 'https://archive.ics.uci.edu/ml/datasets/Chess+%28King-Rook+vs.+King%2
9',
 'https://archive.ics.uci.edu/ml/datasets/Chess+%28Domain+Theories%29',
 'https://archive.ics.uci.edu/ml/datasets/Bach+Chorales',
 'https://archive.ics.uci.edu/ml/datasets/Connect-4',
 'https://archive.ics.uci.edu/ml/datasets/Credit+Approval',
 'https://archive.ics.uci.edu/ml/datasets/Japanese+Credit+Screening',
 'https://archive.ics.uci.edu/ml/datasets/Computer+Hardware',
 'https://archive.ics.uci.edu/ml/datasets/Contraceptive+Method+Choice',
 'https://archive.ics.uci.edu/ml/datasets/Covertype',
 'https://archive.ics.uci.edu/ml/datasets/Cylinder+Bands',
 'https://archive.ics.uci.edu/ml/datasets/Dermatology',
 'https://archive.ics.uci.edu/ml/datasets/Diabetes',
 'https://archive.ics.uci.edu/ml/datasets/DGP2+-+The+Second+Data+Genera
tion+Program',
 'https://archive.ics.uci.edu/ml/datasets/Document+Understanding',
 'https://archive.ics.uci.edu/ml/datasets/EBL+Domain+Theories',
 'https://archive.ics.uci.edu/ml/datasets/Echocardiogram',
 'https://archive.ics.uci.edu/ml/datasets/Ecoli',
 'https://archive.ics.uci.edu/ml/datasets/Flags',
```

```
 'https://archive.ics.uci.edu/ml/datasets/Function+Finding',
 'https://archive.ics.uci.edu/ml/datasets/Glass+Identification',
 'https://archive.ics.uci.edu/ml/datasets/Haberman%27s+Survival',
 'https://archive.ics.uci.edu/ml/datasets/Hayes-Roth',
 'https://archive.ics.uci.edu/ml/datasets/Heart+Disease',
 'https://archive.ics.uci.edu/ml/datasets/Hepatitis',
 'https://archive.ics.uci.edu/ml/datasets/Horse+Colic',
 'https://archive.ics.uci.edu/ml/datasets/ICU',
 'https://archive.ics.uci.edu/ml/datasets/Image+Segmentation',
 'https://archive.ics.uci.edu/ml/datasets/Internet+Advertisements',
 'https://archive.ics.uci.edu/ml/datasets/Ionosphere',
 'https://archive.ics.uci.edu/ml/datasets/Iris',
 'https://archive.ics.uci.edu/ml/datasets/ISOLET',
 'https://archive.ics.uci.edu/ml/datasets/Kinship',
 'https://archive.ics.uci.edu/ml/datasets/Labor+Relations',
 'https://archive.ics.uci.edu/ml/datasets/LED+Display+Domain',
 'https://archive.ics.uci.edu/ml/datasets/Lenses',
 'https://archive.ics.uci.edu/ml/datasets/Letter+Recognition',
 'https://archive.ics.uci.edu/ml/datasets/Liver+Disorders',
 'https://archive.ics.uci.edu/ml/datasets/Logic+Theorist',
 'https://archive.ics.uci.edu/ml/datasets/Lung+Cancer',
 'https://archive.ics.uci.edu/ml/datasets/Lymphography',
 'https://archive.ics.uci.edu/ml/datasets/Mechanical+Analysis',
 'https://archive.ics.uci.edu/ml/datasets/Meta-data',
 'https://archive.ics.uci.edu/ml/datasets/Mobile+Robots',
 'https://archive.ics.uci.edu/ml/datasets/Molecular+Biology+%28Promoter
+Gene+Sequences%29',
 'https://archive.ics.uci.edu/ml/datasets/Molecular+Biology+%28Protein+
Secondary+Structure%29',
 'https://archive.ics.uci.edu/ml/datasets/Molecular+Biology+%28Splice-j
unction+Gene+Sequences%29',
 'https://archive.ics.uci.edu/ml/datasets/MONK%27s+Problems',
 'https://archive.ics.uci.edu/ml/datasets/Moral+Reasoner',
 'https://archive.ics.uci.edu/ml/datasets/Multiple+Features',
 'https://archive.ics.uci.edu/ml/datasets/Mushroom',
 'https://archive.ics.uci.edu/ml/datasets/Musk+%28Version+1%29',
 'https://archive.ics.uci.edu/ml/datasets/Musk+%28Version+2%29',
 'https://archive.ics.uci.edu/ml/datasets/Nursery',
 'https://archive.ics.uci.edu/ml/datasets/Othello+Domain+Theory',
```

```
 'https://archive.ics.uci.edu/ml/datasets/Page+Blocks+Classification',
 'https://archive.ics.uci.edu/ml/datasets/Optical+Recognition+of+Handwr
itten+Digits',
 'https://archive.ics.uci.edu/ml/datasets/Pen-Based+Recognition+of+Hand
written+Digits',
 'https://archive.ics.uci.edu/ml/datasets/Post-Operative+Patient',
 'https://archive.ics.uci.edu/ml/datasets/Primary+Tumor',
 'https://archive.ics.uci.edu/ml/datasets/Prodigy',
 'https://archive.ics.uci.edu/ml/datasets/Qualitative+Structure+Activit
y+Relationships',
 'https://archive.ics.uci.edu/ml/datasets/Quadruped+Mammals',
 'https://archive.ics.uci.edu/ml/datasets/Servo',
 'https://archive.ics.uci.edu/ml/datasets/Shuttle+Landing+Control',
 'https://archive.ics.uci.edu/ml/datasets/Solar+Flare',
 'https://archive.ics.uci.edu/ml/datasets/Soybean+%28Large%29',
 'https://archive.ics.uci.edu/ml/datasets/Soybean+%28Small%29',
 'https://archive.ics.uci.edu/ml/datasets/Challenger+USA+Space+Shuttle+
O-Ring',
 'https://archive.ics.uci.edu/ml/datasets/Low+Resolution+Spectrometer',
 'https://archive.ics.uci.edu/ml/datasets/Spambase',
 'https://archive.ics.uci.edu/ml/datasets/SPECT+Heart',
 'https://archive.ics.uci.edu/ml/datasets/SPECTF+Heart',
 'https://archive.ics.uci.edu/ml/datasets/Sponge',
 'https://archive.ics.uci.edu/ml/datasets/Statlog+Project',
 'https://archive.ics.uci.edu/ml/datasets/Student+Loan+Relational',
 'https://archive.ics.uci.edu/ml/datasets/Teaching+Assistant+Evaluatio
n',
 'https://archive.ics.uci.edu/ml/datasets/Tic-Tac-Toe+Endgame',
 'https://archive.ics.uci.edu/ml/datasets/Thyroid+Disease',
 'https://archive.ics.uci.edu/ml/datasets/Trains',
 'https://archive.ics.uci.edu/ml/datasets/University',
 'https://archive.ics.uci.edu/ml/datasets/Congressional+Voting+Record
s',
 'https://archive.ics.uci.edu/ml/datasets/Water+Treatment+Plant',
 'https://archive.ics.uci.edu/ml/datasets/Waveform+Database+Generator+%
28Version+1%29',
 'https://archive.ics.uci.edu/ml/datasets/Waveform+Database+Generator+%
28Version+2%29',
 'https://archive.ics.uci.edu/ml/datasets/Wine',
```

```
 'https://archive.ics.uci.edu/ml/datasets/Yeast',
 'https://archive.ics.uci.edu/ml/datasets/Zoo',
 'https://archive.ics.uci.edu/ml/datasets/Undocumented',
 'https://archive.ics.uci.edu/ml/datasets/Twenty+Newsgroups',
 'https://archive.ics.uci.edu/ml/datasets/Australian+Sign+Language+sign
s',
 'https://archive.ics.uci.edu/ml/datasets/Australian+Sign+Language+sign
s+%28High+Quality%29',
 'https://archive.ics.uci.edu/ml/datasets/US+Census+Data+%281990%29',
 'https://archive.ics.uci.edu/ml/datasets/Census-Income+%28KDD%29',
 'https://archive.ics.uci.edu/ml/datasets/Coil+1999+Competition+Data',
 'https://archive.ics.uci.edu/ml/datasets/Corel+Image+Features',
 'https://archive.ics.uci.edu/ml/datasets/E.+Coli+Genes',
 'https://archive.ics.uci.edu/ml/datasets/EEG+Database',
 'https://archive.ics.uci.edu/ml/datasets/El+Nino',
 'https://archive.ics.uci.edu/ml/datasets/Entree+Chicago+Recommendation
+Data',
 'https://archive.ics.uci.edu/ml/datasets/CMU+Face+Images',
 'https://archive.ics.uci.edu/ml/datasets/Insurance+Company+Benchmark+%
28COIL+2000%29',
 'https://archive.ics.uci.edu/ml/datasets/Internet+Usage+Data',
 'https://archive.ics.uci.edu/ml/datasets/IPUMS+Census+Database',
 'https://archive.ics.uci.edu/ml/datasets/Japanese+Vowels',
 'https://archive.ics.uci.edu/ml/datasets/KDD+Cup+1998+Data',
 'https://archive.ics.uci.edu/ml/datasets/KDD+Cup+1999+Data',
 'https://archive.ics.uci.edu/ml/datasets/M.+Tuberculosis+Genes',
 'https://archive.ics.uci.edu/ml/datasets/Movie',
 'https://archive.ics.uci.edu/ml/datasets/MSNBC.com+Anonymous+Web+Dat
a',
 'https://archive.ics.uci.edu/ml/datasets/NSF+Research+Award+Abstracts+
1990-2003',
 'https://archive.ics.uci.edu/ml/datasets/Pioneer-1+Mobile+Robot+Data',
 'https://archive.ics.uci.edu/ml/datasets/Pseudo+Periodic+Synthetic+Tim
e+Series',
 'https://archive.ics.uci.edu/ml/datasets/Reuters-21578+Text+Categoriza
tion+Collection',
 'https://archive.ics.uci.edu/ml/datasets/Robot+Execution+Failures',
 'https://archive.ics.uci.edu/ml/datasets/Synthetic+Control+Chart+Time+
Series',
```

```
 'https://archive.ics.uci.edu/ml/datasets/Syskill+and+Webert+Web+Page+R
atings',
 'https://archive.ics.uci.edu/ml/datasets/UNIX+User+Data',
 'https://archive.ics.uci.edu/ml/datasets/Volcanoes+on+Venus+-+JARtool+
experiment',
 'https://archive.ics.uci.edu/ml/datasets/Statlog+%28Australian+Credit+
Approval%29',
 'https://archive.ics.uci.edu/ml/datasets/Statlog+%28German+Credit+Dat
a%29',
 'https://archive.ics.uci.edu/ml/datasets/Statlog+%28Heart%29',
 'https://archive.ics.uci.edu/ml/datasets/Statlog+%28Landsat+Satellite%
29',
 'https://archive.ics.uci.edu/ml/datasets/Statlog+%28Image+Segmentatio
n%29',
 'https://archive.ics.uci.edu/ml/datasets/Statlog+%28Shuttle%29',
 'https://archive.ics.uci.edu/ml/datasets/Statlog+%28Vehicle+Silhouette
s%29',
 'https://archive.ics.uci.edu/ml/datasets/Connectionist+Bench+%28Nettal
k+Corpus%29',
 'https://archive.ics.uci.edu/ml/datasets/Connectionist+Bench+%28Sonar%
2C+Mines+vs.+Rocks%29',
 'https://archive.ics.uci.edu/ml/datasets/Connectionist+Bench+%28Vowel+
Recognition+-+Deterding+Data%29',
 'https://archive.ics.uci.edu/ml/datasets/Economic+Sanctions',
 'https://archive.ics.uci.edu/ml/datasets/Protein+Data',
 'https://archive.ics.uci.edu/ml/datasets/Cloud',
 'https://archive.ics.uci.edu/ml/datasets/CalIt2+Building+People+Count
s',
 'https://archive.ics.uci.edu/ml/datasets/Dodgers+Loop+Sensor',
 'https://archive.ics.uci.edu/ml/datasets/Poker+Hand',
 'https://archive.ics.uci.edu/ml/datasets/MAGIC+Gamma+Telescope',
 'https://archive.ics.uci.edu/ml/datasets/UJI+Pen+Characters',
 'https://archive.ics.uci.edu/ml/datasets/Mammographic+Mass',
 'https://archive.ics.uci.edu/ml/datasets/Forest+Fires',
 'https://archive.ics.uci.edu/ml/datasets/Reuters+Transcribed+Subset',
 'https://archive.ics.uci.edu/ml/datasets/Bag+of+Words',
 'https://archive.ics.uci.edu/ml/datasets/Concrete+Compressive+Strengt
h',
 'https://archive.ics.uci.edu/ml/datasets/Hill-Valley',
```

```
 'https://archive.ics.uci.edu/ml/datasets/Arcene',
 'https://archive.ics.uci.edu/ml/datasets/Dexter',
 'https://archive.ics.uci.edu/ml/datasets/Dorothea',
 'https://archive.ics.uci.edu/ml/datasets/Gisette',
 'https://archive.ics.uci.edu/ml/datasets/Madelon',
 'https://archive.ics.uci.edu/ml/datasets/Ozone+Level+Detection',
 'https://archive.ics.uci.edu/ml/datasets/Abscisic+Acid+Signaling+Netwo
rk',
 'https://archive.ics.uci.edu/ml/datasets/Parkinsons',
 'https://archive.ics.uci.edu/ml/datasets/Character+Trajectories',
 'https://archive.ics.uci.edu/ml/datasets/Blood+Transfusion+Service+Cen
ter',
 'https://archive.ics.uci.edu/ml/datasets/UJI+Pen+Characters+%28Version
+2%29',
 'https://archive.ics.uci.edu/ml/datasets/Semeion+Handwritten+Digit',
 'https://archive.ics.uci.edu/ml/datasets/SECOM',
 'https://archive.ics.uci.edu/ml/datasets/Plants',
 'https://archive.ics.uci.edu/ml/datasets/Libras+Movement',
 'https://archive.ics.uci.edu/ml/datasets/Concrete+Slump+Test',
 'https://archive.ics.uci.edu/ml/datasets/Communities+and+Crime',
 'https://archive.ics.uci.edu/ml/datasets/Acute+Inflammations',
 'https://archive.ics.uci.edu/ml/datasets/Wine+Quality',
 'https://archive.ics.uci.edu/ml/datasets/URL+Reputation',
 'https://archive.ics.uci.edu/ml/datasets/p53+Mutants',
 'https://archive.ics.uci.edu/ml/datasets/Parkinsons+Telemonitoring',
 'https://archive.ics.uci.edu/ml/datasets/Demospongiae',
 'https://archive.ics.uci.edu/ml/datasets/Opinosis+Opinion+%26frasl%3B+
Review',
 'https://archive.ics.uci.edu/ml/datasets/Breast+Tissue',
 'https://archive.ics.uci.edu/ml/datasets/Cardiotocography',
 'https://archive.ics.uci.edu/ml/datasets/Wall-Following+Robot+Navigati
on+Data',
 'https://archive.ics.uci.edu/ml/datasets/Spoken+Arabic+Digit',
 'https://archive.ics.uci.edu/ml/datasets/Localization+Data+for+Person+
Activity',
 'https://archive.ics.uci.edu/ml/datasets/AutoUniv',
 'https://archive.ics.uci.edu/ml/datasets/Steel+Plates+Faults',
 'https://archive.ics.uci.edu/ml/datasets/MiniBooNE+particle+identifica
tion',
```

```
 'https://archive.ics.uci.edu/ml/datasets/YearPredictionMSD',
 'https://archive.ics.uci.edu/ml/datasets/PEMS-SF',
 'https://archive.ics.uci.edu/ml/datasets/OpinRank+Review+Dataset',
 'https://archive.ics.uci.edu/ml/datasets/Relative+location+of+CT+slice
s+on+axial+axis',
 'https://archive.ics.uci.edu/ml/datasets/Online+Handwritten+Assamese+C
haracters+Dataset',
 'https://archive.ics.uci.edu/ml/datasets/PubChem+Bioassay+Data',
 'https://archive.ics.uci.edu/ml/datasets/Record+Linkage+Comparison+Pat
terns',
 'https://archive.ics.uci.edu/ml/datasets/Communities+and+Crime+Unnorma
lized',
 'https://archive.ics.uci.edu/ml/datasets/Vertebral+Column',
 'https://archive.ics.uci.edu/ml/datasets/EMG+Physical+Action+Data+Se
t',
 'https://archive.ics.uci.edu/ml/datasets/Vicon+Physical+Action+Data+Se
t',
 'https://archive.ics.uci.edu/ml/datasets/Amazon+Commerce+reviews+set',
 'https://archive.ics.uci.edu/ml/datasets/Amazon+Access+Samples',
 'https://archive.ics.uci.edu/ml/datasets/Reuter_50_50',
 'https://archive.ics.uci.edu/ml/datasets/Farm+Ads',
 'https://archive.ics.uci.edu/ml/datasets/DBWorld+e-mails',
 'https://archive.ics.uci.edu/ml/datasets/KEGG+Metabolic+Relation+Netwo
rk+%28Directed%29',
 'https://archive.ics.uci.edu/ml/datasets/KEGG+Metabolic+Reaction+Netwo
rk+%28Undirected%29',
 'https://archive.ics.uci.edu/ml/datasets/Bank+Marketing',
 'https://archive.ics.uci.edu/ml/datasets/YouTube+Comedy+Slam+Preferenc
e+Data',
 'https://archive.ics.uci.edu/ml/datasets/Gas+Sensor+Array+Drift+Datase
t',
 'https://archive.ics.uci.edu/ml/datasets/ILPD+%28Indian+Liver+Patient+
Dataset%29',
 'https://archive.ics.uci.edu/ml/datasets/OPPORTUNITY+Activity+Recognit
ion',
 'https://archive.ics.uci.edu/ml/datasets/Nomao',
 'https://archive.ics.uci.edu/ml/datasets/SMS+Spam+Collection',
 'https://archive.ics.uci.edu/ml/datasets/Skin+Segmentation',
 'https://archive.ics.uci.edu/ml/datasets/Planning+Relax',
```

```
 'https://archive.ics.uci.edu/ml/datasets/PAMAP2+Physical+Activity+Moni
toring',
 'https://archive.ics.uci.edu/ml/datasets/Restaurant+%26+consumer+dat
a',
 'https://archive.ics.uci.edu/ml/datasets/CNAE-9',
 'https://archive.ics.uci.edu/ml/datasets/Individual+household+electric
+power+consumption',
 'https://archive.ics.uci.edu/ml/datasets/seeds',
 'https://archive.ics.uci.edu/ml/datasets/Northix',
 'https://archive.ics.uci.edu/ml/datasets/QtyT40I10D100K',
 'https://archive.ics.uci.edu/ml/datasets/Legal+Case+Reports',
 'https://archive.ics.uci.edu/ml/datasets/Human+Activity+Recognition+Us
ing+Smartphones',
 'https://archive.ics.uci.edu/ml/datasets/One-hundred+plant+species+lea
ves+data+set',
 'https://archive.ics.uci.edu/ml/datasets/Energy+efficiency',
 'https://archive.ics.uci.edu/ml/datasets/Yacht+Hydrodynamics',
 'https://archive.ics.uci.edu/ml/datasets/Fertility',
 'https://archive.ics.uci.edu/ml/datasets/Daphnet+Freezing+of+Gait',
 'https://archive.ics.uci.edu/ml/datasets/3D+Road+Network+%28North+Jutl
and%2C+Denmark%29',
 'https://archive.ics.uci.edu/ml/datasets/ISTANBUL+STOCK+EXCHANGE',
 'https://archive.ics.uci.edu/ml/datasets/Buzz+in+social+media+',
 'https://archive.ics.uci.edu/ml/datasets/First-order+theorem+proving',
 'https://archive.ics.uci.edu/ml/datasets/Wearable+Computing%3A+Classif
ication+of+Body+Postures+and+Movements+%28PUC-Rio%29',
 'https://archive.ics.uci.edu/ml/datasets/Gas+sensor+arrays+in+open+sam
pling+settings',
 'https://archive.ics.uci.edu/ml/datasets/Climate+Model+Simulation+Cras
hes',
 'https://archive.ics.uci.edu/ml/datasets/MicroMass',
 'https://archive.ics.uci.edu/ml/datasets/QSAR+biodegradation',
 'https://archive.ics.uci.edu/ml/datasets/BLOGGER',
 'https://archive.ics.uci.edu/ml/datasets/Daily+and+Sports+Activities',
 'https://archive.ics.uci.edu/ml/datasets/User+Knowledge+Modeling',
 'https://archive.ics.uci.edu/ml/datasets/Reuters+RCV1+RCV2+Multilingua
l%2C+Multiview+Text+Categorization+Test+collection',
 'https://archive.ics.uci.edu/ml/datasets/NYSK',
 'https://archive.ics.uci.edu/ml/datasets/Turkiye+Student+Evaluation',
```

```
 'https://archive.ics.uci.edu/ml/datasets/ser+Knowledge+Modeling+Data+%
28Students%27+Knowledge+Levels+on+DC+Electrical+Machines%29',
 'https://archive.ics.uci.edu/ml/datasets/EEG+Eye+State',
 'https://archive.ics.uci.edu/ml/datasets/Physicochemical+Properties+of
+Protein+Tertiary+Structure',
 'https://archive.ics.uci.edu/ml/datasets/seismic-bumps',
 'https://archive.ics.uci.edu/ml/datasets/banknote+authentication',
 'https://archive.ics.uci.edu/ml/datasets/USPTO+Algorithm+Challenge%2C+
run+by+NASA-Harvard+Tournament+Lab+and+TopCoder++++Problem%3A+Pat',
 'https://archive.ics.uci.edu/ml/datasets/YouTube+Multiview+Video+Games
+Dataset',
 'https://archive.ics.uci.edu/ml/datasets/Gas+Sensor+Array+Drift+Datase
t+at+Different+Concentrations',
 'https://archive.ics.uci.edu/ml/datasets/Activities+of+Daily+Living+%2
8ADLs%29+Recognition+Using+Binary+Sensors',
 'https://archive.ics.uci.edu/ml/datasets/SkillCraft1+Master+Table+Data
set',
 'https://archive.ics.uci.edu/ml/datasets/Weight+Lifting+Exercises+moni
tored+with+Inertial+Measurement+Units',
 'https://archive.ics.uci.edu/ml/datasets/SML2010',
 'https://archive.ics.uci.edu/ml/datasets/Bike+Sharing+Dataset',
 'https://archive.ics.uci.edu/ml/datasets/Predict+keywords+activities+i
n+a+online+social+media',
 'https://archive.ics.uci.edu/ml/datasets/Thoracic+Surgery+Data',
 'https://archive.ics.uci.edu/ml/datasets/EMG+dataset+in+Lower+Limb',
 'https://archive.ics.uci.edu/ml/datasets/SUSY',
 'https://archive.ics.uci.edu/ml/datasets/HIGGS',
 'https://archive.ics.uci.edu/ml/datasets/Qualitative_Bankruptcy',
 'https://archive.ics.uci.edu/ml/datasets/LSVT+Voice+Rehabilitation',
 'https://archive.ics.uci.edu/ml/datasets/Dataset+for+ADL+Recognition+w
ith+Wrist-worn+Accelerometer',
 'https://archive.ics.uci.edu/ml/datasets/Wilt',
 'https://archive.ics.uci.edu/ml/datasets/User+Identification+From+Walk
ing+Activity',
 'https://archive.ics.uci.edu/ml/datasets/Activity+Recognition+from+Sin
gle+Chest-Mounted+Accelerometer',
 'https://archive.ics.uci.edu/ml/datasets/Leaf',
 'https://archive.ics.uci.edu/ml/datasets/Dresses_Attribute_Sales',
 'https://archive.ics.uci.edu/ml/datasets/Tamilnadu+Electricity+Board+H
```

```
ourly+Readings',
 'https://archive.ics.uci.edu/ml/datasets/Airfoil+Self-Noise',
 'https://archive.ics.uci.edu/ml/datasets/Wholesale+customers',
 'https://archive.ics.uci.edu/ml/datasets/Twitter+Data+set+for+Arabic+S
entiment+Analysis',
 'https://archive.ics.uci.edu/ml/datasets/Combined+Cycle+Power+Plant',
 'https://archive.ics.uci.edu/ml/datasets/Urban+Land+Cover',
 'https://archive.ics.uci.edu/ml/datasets/Diabetes+130-US+hospitals+for
+years+1999-2008',
 'https://archive.ics.uci.edu/ml/datasets/Bach+Choral+Harmony',
 'https://archive.ics.uci.edu/ml/datasets/StoneFlakes',
 'https://archive.ics.uci.edu/ml/datasets/Tennis+Major+Tournament+Match
+Statistics',
 'https://archive.ics.uci.edu/ml/datasets/Parkinson+Speech+Dataset+with
++Multiple+Types+of+Sound+Recordings',
 'https://archive.ics.uci.edu/ml/datasets/Gesture+Phase+Segmentation',
 'https://archive.ics.uci.edu/ml/datasets/Perfume+Data',
 'https://archive.ics.uci.edu/ml/datasets/BlogFeedback',
 'https://archive.ics.uci.edu/ml/datasets/REALDISP+Activity+Recognition
+Dataset',
 'https://archive.ics.uci.edu/ml/datasets/Newspaper+and+magazine+images
+segmentation+dataset',
 'https://archive.ics.uci.edu/ml/datasets/AAAI+2014+Accepted+Papers',
 'https://archive.ics.uci.edu/ml/datasets/Gas+sensor+array+under+flow+m
odulation',
 'https://archive.ics.uci.edu/ml/datasets/Gas+sensor+array+exposed+to+t
urbulent+gas+mixtures',
 'https://archive.ics.uci.edu/ml/datasets/UJIIndoorLoc',
 'https://archive.ics.uci.edu/ml/datasets/Sentence+Classification',
 'https://archive.ics.uci.edu/ml/datasets/Dow+Jones+Index',
 'https://archive.ics.uci.edu/ml/datasets/sEMG+for+Basic+Hand+movement
s',
 'https://archive.ics.uci.edu/ml/datasets/AAAI+2013+Accepted+Papers',
 'https://archive.ics.uci.edu/ml/datasets/Geographical+Original+of+Musi
c',
 'https://archive.ics.uci.edu/ml/datasets/Condition+Based+Maintenance+o
f+Naval+Propulsion+Plants',
 'https://archive.ics.uci.edu/ml/datasets/Grammatical+Facial+Expression
s',
```

```
 'https://archive.ics.uci.edu/ml/datasets/NoisyOffice',
 'https://archive.ics.uci.edu/ml/datasets/MHEALTH+Dataset',
 'https://archive.ics.uci.edu/ml/datasets/Student+Performance',
 'https://archive.ics.uci.edu/ml/datasets/ElectricityLoadDiagrams201120
14',
 'https://archive.ics.uci.edu/ml/datasets/Gas+sensor+array+under+dynami
c+gas+mixtures',
 'https://archive.ics.uci.edu/ml/datasets/microblogPCU',
 'https://archive.ics.uci.edu/ml/datasets/Firm-Teacher_Clave-Direction_
Classification',
 'https://archive.ics.uci.edu/ml/datasets/Dataset+for+Sensorless+Drive+
Diagnosis',
 'https://archive.ics.uci.edu/ml/datasets/TV+News+Channel+Commercial+De
tection+Dataset',
 'https://archive.ics.uci.edu/ml/datasets/Phishing+Websites',
 'https://archive.ics.uci.edu/ml/datasets/Greenhouse+Gas+Observing+Netw
ork',
 'https://archive.ics.uci.edu/ml/datasets/Diabetic+Retinopathy+Debrecen
+Data+Set',
 'https://archive.ics.uci.edu/ml/datasets/HIV-1+protease+cleavage',
 'https://archive.ics.uci.edu/ml/datasets/Sentiment+Labelled+Sentence
s',
 'https://archive.ics.uci.edu/ml/datasets/Online+News+Popularity',
 'https://archive.ics.uci.edu/ml/datasets/Forest+type+mapping',
 'https://archive.ics.uci.edu/ml/datasets/wiki4HE',
 'https://archive.ics.uci.edu/ml/datasets/Online+Video+Characteristics+
and+Transcoding+Time+Dataset',
 'https://archive.ics.uci.edu/ml/datasets/Chronic_Kidney_Disease',
 'https://archive.ics.uci.edu/ml/datasets/Machine+Learning+based+ZZAlph
a+Ltd.+Stock+Recommendations+2012-2014',
 'https://archive.ics.uci.edu/ml/datasets/Folio',
 'https://archive.ics.uci.edu/ml/datasets/Taxi+Service+Trajectory+-+Pre
diction+Challenge%2C+ECML+PKDD+2015',
 'https://archive.ics.uci.edu/ml/datasets/Cuff-Less+Blood+Pressure+Esti
mation',
 'https://archive.ics.uci.edu/ml/datasets/Smartphone-Based+Recognition+
of+Human+Activities+and+Postural+Transitions',
 'https://archive.ics.uci.edu/ml/datasets/Mice+Protein+Expression',
 'https://archive.ics.uci.edu/ml/datasets/UJIIndoorLoc-Mag',
```

```
 'https://archive.ics.uci.edu/ml/datasets/Heterogeneity+Activity+Recogn
ition',
 'https://archive.ics.uci.edu/ml/datasets/Educational+Process+Mining+%2
8EPM%29%3A+A+Learning+Analytics+Data+Set',
 'https://archive.ics.uci.edu/ml/datasets/HEPMASS',
 'https://archive.ics.uci.edu/ml/datasets/Indoor+User+Movement+Predicti
on+from+RSS+data',
 'https://archive.ics.uci.edu/ml/datasets/Open+University+Learning+Anal
ytics+dataset',
 'https://archive.ics.uci.edu/ml/datasets/default+of+credit+card+client
s',
 'https://archive.ics.uci.edu/ml/datasets/Mesothelioma%E2%80%99s+diseas
e+data+set+',
 'https://archive.ics.uci.edu/ml/datasets/Online+Retail',
 'https://archive.ics.uci.edu/ml/datasets/SIFT10M',
 'https://archive.ics.uci.edu/ml/datasets/GPS+Trajectories',
 'https://archive.ics.uci.edu/ml/datasets/Detect+Malacious+Executable%2
8AntiVirus%29',
 'https://archive.ics.uci.edu/ml/datasets/Occupancy+Detection+',
 'https://archive.ics.uci.edu/ml/datasets/Improved+Spiral+Test+Using+Di
gitized+Graphics+Tablet+for+Monitoring+Parkinson%E2%80%99s+Disease',
 'https://archive.ics.uci.edu/ml/datasets/News+Aggregator',
 'https://archive.ics.uci.edu/ml/datasets/Air+Quality',
 'https://archive.ics.uci.edu/ml/datasets/Twin+gas+sensor+arrays',
 'https://archive.ics.uci.edu/ml/datasets/Gas+sensors+for+home+activity
+monitoring',
 'https://archive.ics.uci.edu/ml/datasets/Facebook+Comment+Volume+Datas
et',
 'https://archive.ics.uci.edu/ml/datasets/Smartphone+Dataset+for+Human+
Activity+Recognition+%28HAR%29+in+Ambient+Assisted+Living+%28AAL%29',
 'https://archive.ics.uci.edu/ml/datasets/Polish+companies+bankruptcy+d
ata',
 'https://archive.ics.uci.edu/ml/datasets/Activity+Recognition+system+b
ased+on+Multisensor+data+fusion+%28AReM%29',
 'https://archive.ics.uci.edu/ml/datasets/Dota2+Games+Results',
 'https://archive.ics.uci.edu/ml/datasets/Facebook+metrics',
 'https://archive.ics.uci.edu/ml/datasets/UbiqLog+%28smartphone+lifelog
ging%29',
 'https://archive.ics.uci.edu/ml/datasets/NIPS+Conference+Papers+1987-2
```

```
015',
 'https://archive.ics.uci.edu/ml/datasets/HTRU2',
 'https://archive.ics.uci.edu/ml/datasets/Drug+consumption+%28quantifie
d%29',
 'https://archive.ics.uci.edu/ml/datasets/Appliances+energy+predictio
n',
 'https://archive.ics.uci.edu/ml/datasets/Miskolc+IIS+Hybrid+IPS',
 'https://archive.ics.uci.edu/ml/datasets/KDC-4007+dataset+Collection',
 'https://archive.ics.uci.edu/ml/datasets/Geo-Magnetic+field+and+WLAN+d
ataset+for+indoor+localisation+from+wristband+and+smartphone',
 'https://archive.ics.uci.edu/ml/datasets/DrivFace',
 'https://archive.ics.uci.edu/ml/datasets/Website+Phishing',
 'https://archive.ics.uci.edu/ml/datasets/YouTube+Spam+Collection',
 'https://archive.ics.uci.edu/ml/datasets/Beijing+PM2.5+Data',
 'https://archive.ics.uci.edu/ml/datasets/Cargo+2000+Freight+Tracking+a
nd+Tracing',
 'https://archive.ics.uci.edu/ml/datasets/Cervical+cancer+%28Risk+Facto
rs%29',
 'https://archive.ics.uci.edu/ml/datasets/Quality+Assessment+of+Digital
+Colposcopies',
 'https://archive.ics.uci.edu/ml/datasets/KASANDR',
 'https://archive.ics.uci.edu/ml/datasets/FMA%3A+A+Dataset+For+Music+An
alysis',
 'https://archive.ics.uci.edu/ml/datasets/Air+quality',
 'https://archive.ics.uci.edu/ml/datasets/Epileptic+Seizure+Recognitio
n',
 'https://archive.ics.uci.edu/ml/datasets/Devanagari+Handwritten+Charac
ter+Dataset',
 'https://archive.ics.uci.edu/ml/datasets/Stock+portfolio+performance',
 'https://archive.ics.uci.edu/ml/datasets/MoCap+Hand+Postures',
 'https://archive.ics.uci.edu/ml/datasets/Early+biomarkers+of+Parkinso
n%92s+disease+based+on+natural+connected+speech',
 'https://archive.ics.uci.edu/ml/datasets/Data+for+Software+Engineering
+Teamwork+Assessment+in+Education+Setting',
 'https://archive.ics.uci.edu/ml/datasets/PM2.5+Data+of+Five+Chinese+Ci
ties',
 'https://archive.ics.uci.edu/ml/datasets/Parkinson+Disease+Spiral+Draw
ings+Using+Digitized+Graphics+Tablet',
 'https://archive.ics.uci.edu/ml/datasets/Sales_Transactions_Dataset_We
```

```
ekly',
 'https://archive.ics.uci.edu/ml/datasets/Las+Vegas+Strip',
 'https://archive.ics.uci.edu/ml/datasets/Eco-hotel',
 'https://archive.ics.uci.edu/ml/datasets/MEU-Mobile+KSD',
 'https://archive.ics.uci.edu/ml/datasets/Crowdsourced+Mapping',
 'https://archive.ics.uci.edu/ml/datasets/gene+expression+cancer+RNA-Se
q',
 'https://archive.ics.uci.edu/ml/datasets/Hybrid+Indoor+Positioning+Dat
aset+from+WiFi+RSSI%2C+Bluetooth+and+magnetometer',
 'https://archive.ics.uci.edu/ml/datasets/chestnut+%E2%80%93+LARVIC',
 'https://archive.ics.uci.edu/ml/datasets/Burst+Header+Packet+%28BHP%29
+flooding+attack+on+Optical+Burst+Switching+%28OBS%29+Network',
 'https://archive.ics.uci.edu/ml/datasets/Motion+Capture+Hand+Posture
s',
 'https://archive.ics.uci.edu/ml/datasets/Anuran+Calls+%28MFCCs%29',
 'https://archive.ics.uci.edu/ml/datasets/TTC-3600%3A+Benchmark+dataset
+for+Turkish+text+categorization',
 'https://archive.ics.uci.edu/ml/datasets/Gastrointestinal+Lesions+in+R
egular+Colonoscopy',
 'https://archive.ics.uci.edu/ml/datasets/Daily+Demand+Forecasting+Orde
rs',
 'https://archive.ics.uci.edu/ml/datasets/Paper+Reviews',
 'https://archive.ics.uci.edu/ml/datasets/extention+of+Z-Alizadeh+sani+
dataset',
 'https://archive.ics.uci.edu/ml/datasets/Z-Alizadeh+Sani',
 'https://archive.ics.uci.edu/ml/datasets/Dynamic+Features+of+VirusShar
e+Executables',
 'https://archive.ics.uci.edu/ml/datasets/IDA2016Challenge',
 'https://archive.ics.uci.edu/ml/datasets/DSRC+Vehicle+Communications',
 'https://archive.ics.uci.edu/ml/datasets/Mturk+User-Perceived+Clusters
+over+Images',
 'https://archive.ics.uci.edu/ml/datasets/Character+Font+Images',
 'https://archive.ics.uci.edu/ml/datasets/DeliciousMIL%3A+A+Data+Set+fo
r+Multi-Label+Multi-Instance+Learning+with+Instance+Labels',
 'https://archive.ics.uci.edu/ml/datasets/Autistic+Spectrum+Disorder+Sc
reening+Data+for+Children++',
 'https://archive.ics.uci.edu/ml/datasets/Autistic+Spectrum+Disorder+Sc
reening+Data+for+Adolescent+++',
 'https://archive.ics.uci.edu/ml/datasets/APS+Failure+at+Scania+Truck
```

```
s',
 'https://archive.ics.uci.edu/ml/datasets/Wireless+Indoor+Localizatio
n',
 'https://archive.ics.uci.edu/ml/datasets/HCC+Survival',
 'https://archive.ics.uci.edu/ml/datasets/CSM+%28Conventional+and+Socia
l+Media+Movies%29+Dataset+2014+and+2015',
 'https://archive.ics.uci.edu/ml/datasets/University+of+Tehran+Question
+Dataset+2016+%28UTQD.2016%29',
 'https://archive.ics.uci.edu/ml/datasets/Autism+Screening+Adult',
 'https://archive.ics.uci.edu/ml/datasets/Activity+recognition+with+hea
lthy+older+people+using+a+batteryless+wearable+sensor',
 'https://archive.ics.uci.edu/ml/datasets/Immunotherapy+Dataset',
 'https://archive.ics.uci.edu/ml/datasets/Cryotherapy+Dataset+',
 'https://archive.ics.uci.edu/ml/datasets/OCT+data+%26+Color+Fundus+Ima
ges+of+Left+%26+Right+Eyes',
 'https://archive.ics.uci.edu/ml/datasets/Discrete+Tone+Image+Dataset',
 'https://archive.ics.uci.edu/ml/datasets/News+Popularity+in+Multiple+S
ocial+Media+Platforms',
 'https://archive.ics.uci.edu/ml/datasets/Ultrasonic+flowmeter+diagnost
ics',
 'https://archive.ics.uci.edu/ml/datasets/ICMLA+2014+Accepted+Papers+Da
ta+Set',
 'https://archive.ics.uci.edu/ml/datasets/BLE+RSSI+Dataset+for+Indoor+l
ocalization+and+Navigation',
 'https://archive.ics.uci.edu/ml/datasets/Container+Crane+Controller+Da
ta+Set',
 'https://archive.ics.uci.edu/ml/datasets/Residential+Building+Data+Se
t',
 'https://archive.ics.uci.edu/ml/datasets/Health+News+in+Twitter',
 'https://archive.ics.uci.edu/ml/datasets/chipseq',
 'https://archive.ics.uci.edu/ml/datasets/SGEMM+GPU+kernel+performanc
e',
 'https://archive.ics.uci.edu/ml/datasets/Repeat+Consumption+Matrices',
 'https://archive.ics.uci.edu/ml/datasets/detection_of_IoT_botnet_attac
ks_N_BaIoT',
 'https://archive.ics.uci.edu/ml/datasets/Absenteeism+at+work',
 'https://archive.ics.uci.edu/ml/datasets/SCADI',
 'https://archive.ics.uci.edu/ml/datasets/Condition+monitoring+of+hydra
ulic+systems',
```

```
 'https://archive.ics.uci.edu/ml/datasets/Carbon+Nanotubes',
 'https://archive.ics.uci.edu/ml/datasets/Optical+Interconnection+Netwo
rk+',
 'https://archive.ics.uci.edu/ml/datasets/Sports+articles+for+objectivi
ty+analysis',
 'https://archive.ics.uci.edu/ml/datasets/Breast+Cancer+Coimbra',
 'https://archive.ics.uci.edu/ml/datasets/GNFUV+Unmanned+Surface+Vehicl
es+Sensor+Data',
 'https://archive.ics.uci.edu/ml/datasets/Dishonest+Internet+users+Data
set',
 'https://archive.ics.uci.edu/ml/datasets/Victorian+Era+Authorship+Attr
ibution',
 'https://archive.ics.uci.edu/ml/datasets/Simulated+Falls+and+Daily+Liv
ing+Activities+Data+Set',
 'https://archive.ics.uci.edu/ml/datasets/Multimodal+Damage+Identificat
ion+for+Humanitarian+Computing',
 'https://archive.ics.uci.edu/ml/datasets/EEG+Steady-State+Visual+Evoke
d+Potential+Signals',
 'https://archive.ics.uci.edu/ml/datasets/Roman+Urdu+Data+Set',
 'https://archive.ics.uci.edu/ml/datasets/Avila',
 'https://archive.ics.uci.edu/ml/datasets/PANDOR',
 'https://archive.ics.uci.edu/ml/datasets/Drug+Review+Dataset+%28Drugli
b.com%29',
 'https://archive.ics.uci.edu/ml/datasets/Drug+Review+Dataset+%28Drugs.
com%29',
 'https://archive.ics.uci.edu/ml/datasets/Physical+Unclonable+Function
s',
 'https://archive.ics.uci.edu/ml/datasets/Superconductivty+Data',
 'https://archive.ics.uci.edu/ml/datasets/WESAD+%28Wearable+Stress+and+
Affect+Detection%29',
 'https://archive.ics.uci.edu/ml/datasets/GNFUV+Unmanned+Surface+Vehicl
es+Sensor+Data+Set+2',
 'https://archive.ics.uci.edu/ml/datasets/Student+Academics+Performanc
e',
 'https://archive.ics.uci.edu/ml/datasets/Online+Shoppers+Purchasing+In
tention+Dataset',
 'https://archive.ics.uci.edu/ml/datasets/PMU-UD',
 'https://archive.ics.uci.edu/ml/datasets/Parkinson%27s+Disease+Classif
ication',
```

```
 'https://archive.ics.uci.edu/ml/datasets/Electrical+Grid+Stability+Sim
ulated+Data+',
 'https://archive.ics.uci.edu/ml/datasets/Caesarian+Section+Classificat
ion+Dataset',
 'https://archive.ics.uci.edu/ml/datasets/BAUM-1',
 'https://archive.ics.uci.edu/ml/datasets/BAUM-2',
 'https://archive.ics.uci.edu/ml/datasets/Audit+Data',
 'https://archive.ics.uci.edu/ml/datasets/BuddyMove+Data+Set',
 'https://archive.ics.uci.edu/ml/datasets/Real+estate+valuation+data+se
t',
 'https://archive.ics.uci.edu/ml/datasets/Early+biomarkers+of+Parkinso
n%E2%80%99s+disease+based+on+natural+connected+speech+Data+Set+',
 'https://archive.ics.uci.edu/ml/datasets/Somerville+Happiness+Survey',
 'https://archive.ics.uci.edu/ml/datasets/2.4+GHZ+Indoor+Channel+Measur
ements',
 'https://archive.ics.uci.edu/ml/datasets/EMG+data+for+gestures',
 'https://archive.ics.uci.edu/ml/datasets/Parking+Birmingham',
 'https://archive.ics.uci.edu/ml/datasets/Behavior+of+the+urban+traffic
+of+the+city+of+Sao+Paulo+in+Brazil',
 'https://archive.ics.uci.edu/ml/datasets/Travel+Reviews',
 'https://archive.ics.uci.edu/ml/datasets/Tarvel+Review+Ratings',
 'https://archive.ics.uci.edu/ml/datasets/Rice+Leaf+Diseases',
 'https://archive.ics.uci.edu/ml/datasets/Gas+sensor+array+temperature+
modulation',
 'https://archive.ics.uci.edu/ml/datasets/Facebook+Live+Sellers+in+Thai
land',
 'https://archive.ics.uci.edu/ml/datasets/Parkinson+Dataset+with+replic
ated+acoustic+features+',
 'https://archive.ics.uci.edu/ml/datasets/Metro+Interstate+Traffic+Volu
me',
 'https://archive.ics.uci.edu/ml/datasets/Query+Analytics+Workloads+Dat
aset',
 'https://archive.ics.uci.edu/ml/datasets/Wave+Energy+Converters',
 'https://archive.ics.uci.edu/ml/datasets/PPG-DaLiA',
 'https://archive.ics.uci.edu/ml/datasets/Alcohol+QCM+Sensor+Dataset',
 'https://archive.ics.uci.edu/ml/datasets/Divorce+Predictors+data+set',
 'https://archive.ics.uci.edu/ml/datasets/Incident+management+process+e
nriched+event+log',
 'https://archive.ics.uci.edu/ml/datasets/Opinion+Corpus+for+Lebanese+A
```

```
rabic+Reviews+%28OCLAR%29',
 'https://archive.ics.uci.edu/ml/datasets/MEx',
 'https://archive.ics.uci.edu/ml/datasets/Beijing+Multi-Site+Air-Qualit
y+Data',
 'https://archive.ics.uci.edu/ml/datasets/Online+Retail+II',
 'https://archive.ics.uci.edu/ml/datasets/Hepatitis+C+Virus+%28HCV%29+f
or+Egyptian+patients',
 'https://archive.ics.uci.edu/ml/datasets/QSAR+fish+toxicity',
 'https://archive.ics.uci.edu/ml/datasets/QSAR+aquatic+toxicity',
 'https://archive.ics.uci.edu/ml/datasets/Human+Activity+Recognition+fr
om+Continuous+Ambient+Sensor+Data',
 'https://archive.ics.uci.edu/ml/datasets/WISDM+Smartphone+and+Smartwat
ch+Activity+and+Biometrics+Dataset+',
 'https://archive.ics.uci.edu/ml/datasets/QSAR+oral+toxicity',
 'https://archive.ics.uci.edu/ml/datasets/QSAR+androgen+receptor',
 'https://archive.ics.uci.edu/ml/datasets/QSAR+Bioconcentration+classes
+dataset',
 'https://archive.ics.uci.edu/ml/datasets/QSAR+fish+bioconcentration+fa
ctor+%28BCF%29',
 'https://archive.ics.uci.edu/ml/datasets/A+study+of++Asian+Religious+a
nd+Biblical+Texts',
 'https://archive.ics.uci.edu/ml/datasets/Real-time+Election+Results%3A
+Portugal+2019',
 'https://archive.ics.uci.edu/ml/datasets/Bias+correction+of+numerical+
prediction+model+temperature+forecast',
 'https://archive.ics.uci.edu/ml/datasets/Bar+Crawl%3A+Detecting+Heavy+
Drinking',
 'https://archive.ics.uci.edu/ml/datasets/Kitsune+Network+Attack+Datase
t',
 'https://archive.ics.uci.edu/ml/datasets/Shoulder+Implant+X-Ray+Manufa
cturer+Classification',
 'https://archive.ics.uci.edu/ml/datasets/Speaker+Accent+Recognition',
 'https://archive.ics.uci.edu/ml/datasets/Heart+failure+clinical+record
s',
 'https://archive.ics.uci.edu/ml/datasets/Deepfakes%3A+Medical+Image+Ta
mper+Detection',
 'https://archive.ics.uci.edu/ml/datasets/selfBACK',
 'https://archive.ics.uci.edu/ml/datasets/South+German+Credit',
 'https://archive.ics.uci.edu/ml/datasets/Exasens',
```

```
 'https://archive.ics.uci.edu/ml/datasets/Swarm+Behaviour',
 'https://archive.ics.uci.edu/ml/datasets/Crop+mapping+using+fused+opti
cal-radar+data+set',
 'https://archive.ics.uci.edu/ml/datasets/BitcoinHeistRansomwareAddress
Dataset',
 'https://archive.ics.uci.edu/ml/datasets/Facebook+Large+Page-Page+Netw
ork',
 'https://archive.ics.uci.edu/ml/datasets/Amphibians',
 'https://archive.ics.uci.edu/ml/datasets/Early+stage+diabetes+risk+pre
diction+dataset.',
 'https://archive.ics.uci.edu/ml/datasets/Turkish+Spam+V01',
 'https://archive.ics.uci.edu/ml/datasets/Stock+keeping+units',
 'https://archive.ics.uci.edu/ml/datasets/Demand+Forecasting+for+a+stor
e',
 'https://archive.ics.uci.edu/ml/datasets/Detect+Malware+Types',
 'https://archive.ics.uci.edu/ml/datasets/Wave+Energy+Converters',
 'https://archive.ics.uci.edu/ml/datasets/Youtube+cookery+channels+view
ers+comments+in+Hinglish',
 'https://archive.ics.uci.edu/ml/datasets/Pedestrian+in+Traffic+Datase
t',
 'https://archive.ics.uci.edu/ml/datasets/Cervical+Cancer+Behavior+Ris
k',
 'https://archive.ics.uci.edu/ml/datasets/Sattriya_Dance_Single_Hand_Ge
stures+Dataset',
 'https://archive.ics.uci.edu/ml/datasets/Divorce+Predictors+data+set',
 'https://archive.ics.uci.edu/ml/datasets/3W+dataset',
 'https://archive.ics.uci.edu/ml/datasets/Malware+static+and+dynamic+fe
atures+VxHeaven+and+Virus+Total',
 'https://archive.ics.uci.edu/ml/datasets/Internet+Firewall+Data',
 'https://archive.ics.uci.edu/ml/datasets/User+Profiling+and+Abusive+La
nguage+Detection+Dataset',
 'https://archive.ics.uci.edu/ml/datasets/Estimation+of+obesity+levels+
based+on+eating+habits+and+physical+condition+',
 'https://archive.ics.uci.edu/ml/datasets/Rice+%28Cammeo+and+Osmancik%2
9',
 'https://archive.ics.uci.edu/ml/datasets/Vehicle+routing+and+schedulin
g+problems',
 'https://archive.ics.uci.edu/ml/datasets/Algerian+Forest+Fires+Dataset
++',
```

```
 'https://archive.ics.uci.edu/ml/datasets/Breath+Metabolomics',
 'https://archive.ics.uci.edu/ml/datasets/Horton+General+Hospital',
 'https://archive.ics.uci.edu/ml/datasets/UrbanGB%2C+urban+road+accidents+coordinates+labelled+by+the+urban+center',
 'https://archive.ics.uci.edu/ml/datasets/Gas+Turbine+CO+and+NOx+Emission+Data+Set',
 'https://archive.ics.uci.edu/ml/datasets/Activity+recognition+using+wearable+physiological+measurements',
 'https://archive.ics.uci.edu/ml/datasets/clickstream+data+for+online+shopping',
 'https://archive.ics.uci.edu/ml/datasets/CNNpred%3A+CNN-based+stock+market+prediction+using+a+diverse+set+of+variables',
 'https://archive.ics.uci.edu/ml/datasets/Apartment+for+rent+classified',
 'https://archive.ics.uci.edu/ml/datasets/%3A+Simulated+Data+set+of+Iraqi+tourism+places',
 'https://archive.ics.uci.edu/ml/datasets/Nasarian+CAD+Dataset',
 'https://archive.ics.uci.edu/ml/datasets/Monolithic+Columns+in+Troad+and+Mysia+Region',
 'https://archive.ics.uci.edu/ml/datasets/Bar+Crawl%3A+Detecting+Heavy+Drinking',
 'https://archive.ics.uci.edu/ml/datasets/Seoul+Bike+Sharing+Demand',
 'https://archive.ics.uci.edu/ml/datasets/Person+Classification+Gait+Data',
 'https://archive.ics.uci.edu/ml/datasets/Shill+Bidding+Dataset',
 'https://archive.ics.uci.edu/ml/datasets/Iranian+Churn+Dataset',
 'https://archive.ics.uci.edu/ml/datasets/Unmanned+Aerial+Vehicle+%28UAV%29+Intrusion+Detection',
 'https://archive.ics.uci.edu/ml/datasets/Bone+marrow+transplant%3A+children',
 'https://archive.ics.uci.edu/ml/datasets/Exasens',
 'https://archive.ics.uci.edu/ml/datasets/COVID-19+Surveillance',
 'https://archive.ics.uci.edu/ml/datasets/Refractive+errors',
 'https://archive.ics.uci.edu/ml/datasets/Shoulder+Implant+X-Ray+Manufacturer+Classification',
 'https://archive.ics.uci.edu/ml/datasets/CLINC150',
 'https://archive.ics.uci.edu/ml/datasets/HCV+data',
 'https://archive.ics.uci.edu/ml/datasets/Taiwanese+Bankruptcy+Prediction',
```

```
 'https://archive.ics.uci.edu/ml/datasets/South+German+Credit+%28UPDAT
E%29',
 'https://archive.ics.uci.edu/ml/datasets/IIWA14-R820-Gazebo-Dataset-10
Trajectories',
 'https://archive.ics.uci.edu/ml/datasets/Guitar+Chords+finger+position
s',
 'https://archive.ics.uci.edu/ml/datasets/Russian+Corpus+of+Biographica
l+Texts',
 'https://archive.ics.uci.edu/ml/datasets/Codon+usage',
 'https://archive.ics.uci.edu/ml/datasets/Intelligent+Media+Acceleromet
er+and+Gyroscope+%28IM-AccGyro%29+Dataset']
```

In [136]:
```
len(link)
```

Out[136]: 559

In [137]:
```
type(nombres)
```

Out[137]: list

In [138]:
```
import pandas as pd
dtafrm = {'Data Types':data_types,'Default Task':default_task,'Attribut
e Types':attribute_types,
          'Num Instances':num_instances,'Num Attributes':num_attributes
, 'Year':year,'Link':link}
halfdf = pd.DataFrame(dtafrm,index =nombres)
halfdf
```

Out[138]:

	Data Types	Default Task	Attribute Types	Num Instances	Num Attributes	Year	
Abalone	Multivariate	Classification	Categorical, Integer, Real	4177	8	1995	https://archive
Adult	Multivariate	Classification	Categorical, Integer	48842	14	1996	https://arch
	Data Types	**Default Task**	**Attribute Types**	**Num Instances**	**Num Attributes**	**Year**	

Annealing	Multivariate	Classification	Categorical, Integer, Real	798	38		https://archive.ic
Anonymous Microsoft Web Data		Recommender-Systems	Categorical	37711	294	1998	https://archive.ic
Arrhythmia	Multivariate	Classification	Categorical, Integer, Real	452	279	1998	https://archive
...	...	...	...	...	...	...	
IIWA14-R820-Gazebo-Dataset-10Trajectories		Regression	Integer			2020	https://archive.i
Guitar Chords finger positions	Text	Classification		2633	5	2020	https://archive
Russian Corpus of Biographical Texts	Text	Classification		200	2	2020	https://archive.
Codon usage	Multivariate	Classification, Clustering		13028	69	2020	https://archive.ic
Intelligent Media Accelerometer and Gyroscope (IM-AccGyro) Dataset	Time-Series	Classification	Real	800	9	2020	https://archiv

559 rows × 7 columns

In [139]:
```
#description
import requests
from bs4 import BeautifulSoup
des= requests.get('https://archive.ics.uci.edu/ml/datasets.php?format=&
```

```
task=&att=&area=&numAtt=&numIns=&type=&sort=nameUp&view=list').text
description= BeautifulSoup(des)
```

In [140]:

```
import requests
descrip = re.findall('''</a></b>: ([\w''  "0-9/?'^.%+®\,;*$:!@#(){}_+–
&-<>/~`±]*)</p>''', str(description))
descrip.insert(10,'This dataset comprises information regarding the ADL
s performed by two users on a daily basis in their own homes.')
descrip.insert(15,'The data was created by a medical expert as a data s
et to test the expert system, which will perform the presumptive diagno
sis of two diseases of the urinary system.')
descrip.insert(28,'This is a dataset of classified for apartments for r
ent in USA.')
descrip.insert(46,'The data set is composed of 60 chorales (5665 event
s) by J.S. Bach (1675-1750). Each event of each chorale is labelled usi
ng 1 among 101 chord labels and described through 14 features.')
descrip.insert(61,"It contains fourteen numerical weather prediction (N
WP)'s meteorological forecast data, two in-situ observations, and five
 geographical auxiliary variables over Seoul, South Korea in the summe
r.")
descrip.insert(66,'In this paper, we look for to recognize the causes o
f users tend to cyber space in Kohkiloye and Boyer Ahmad Province in Ir
an')
descrip.insert(97,'ChIP-seq experiments characterize protein modificati
ons or binding at specific genomic locations in specific samples. The m
achine learning problem in these data is structured binary classificati
on.')
descrip.insert(104,'This is a data set containing 1080 documents of fre
e text business descriptions of Brazilian companies categorized into a
 subset of 9 categories')
descrip.insert(133,'The dataset comprises motion sensor data of 19 dail
y and sports activities each performed by 8 subjects in their own style
 for 5 minutes. Five Xsens MTx units are used on the torso, arms, and l
egs.')
descrip.insert(135,"This dataset contains the annotated readings of 3 a
cceleration sensors at the hip and leg of Parkinson's disease patients
 that experience freezing of gait (FoG) during walking tasks.")
descrip.insert(150,'DEXTER is a text classification problem in a bag-of
-word representation. This is a two-class classification problem with s
```

```
parse continuous input variables. This dataset is one of five datasets
 of the NIPS 2003 feature selection challenge.')
descrip.insert(153,'This data has been prepared to analyze factors rela
ted to readmission as well as other outcomes pertaining to patients wit
h diabetes.')
descrip.insert(155,'Discrete Tone Images(DTI)are available which needs
 to be analyzed in detail. Here, we created this dataset for those who
 do research in DTI. ')
descrip.insert(157,'Participants completed the “Personal Information Fo
rm” and “Divorce Predictors Scale”')
descrip.insert(165,'The DrivFace contains images sequences of subjects
 while driving in real scenarios. It is composed of 606 samples of 640×
480, acquired over different days from 4 drivers with several facial fe
atures.')
descrip.insert(174,'This dataset contains the sign and symptpom data of
 newly diabetic or would be diabetic patient.')
descrip.insert(186,'This data set contains electricity consumption of 3
70 points/clients.')
descrip.insert(214,'A chemical detection platform composed of 14 temper
ature-modulated metal oxide (MOX) gas sensors was exposed during 3 week
s to mixtures of carbon monoxide and humid synthetic air in a gas chamb
er.')
descrip.insert(218,'100 recordings of a sensor array under different co
nditions in a home setting: background, wine and banana presentations.
 The array includes 8 MOX gas sensors, and humidity and temperature sen
sors.')
descrip.insert(223,'Instances in this dataset contain audio features ex
tracted from 1059 wave files. The task associated with the data is to p
redict the geographical origin of music. ')
descrip.insert(225,'GISETTE is a handwritten digit recognition problem.
 The problem is to separate the highly confusible digits 4 and 9. This
 dataset is one of five datasets of the NIPS 2003 feature selection cha
llenge.')
descrip.insert(229,'The dataset has been feed by Android app called Go!
Track. It is available at Goolge Play Store(https://play.google.com/sto
re/apps/details?id=com.go.router).')
descrip.insert(231,'Design an observing network to monitor emissions of
 a greenhouse gas (GHG) in California given time series of synthetic ob
servations and tracers from weather model simulations.')
```

```
descrip.insert(265,'The IM-AccGyro dataset is devised to benchmark tech
niques dealing with human activity recognition based on inertial sensor
s')
descrip.insert(292,'The data set contains 15 classes of 24 instances ea
ch. Each class references to a hand movement type in LIBRAS (Portuguese
 name L@ngua BRAsileira de Sinais, oficial brazilian signal language).'
)
descrip.insert(307,'Mesothelioma’s disease data set were prepared at Di
cle University Faculty of Medicine in Turkey. Three hundred and twenty-
four Mesothelioma patient data. In the dataset, all samples have 34 fea
tures.')
descrip.insert(311,'The MEx Multi-modal Exercise dataset contains data
 of 7 different physiotherapy exercises, performed by 30 subjects recor
ded with 2 accelerometers, a pressure mat and a depth camera.')
descrip.insert(314,'MicroblogPCU data is crawled from sina weibo microb
log[http://weibo.com/]. This data can be used to study machine learning
 methods as well as do some social network research.')
descrip.insert(341,'Nomao collects data about places (name, phone, loca
lization...) from many sources. Deduplication consists in detecting wha
t data refer to the same place. Instances in the dataset compare 2 spot
s.')
descrip.insert(349,'This is a dataset of 8235 online handwritten assame
se characters. The “online” process involves capturing of data as text
 is written on a digitizing tablet with an electronic pen.')
descrip.insert(353,'Of the 12,330 sessions in the dataset, 84.5% (10,42
2) were negative class samples that did not end with shopping, and the
 rest (1908) were positive class samples ending with shopping.')
descrip.insert(357,'This dataset contains sentences extracted from user
 reviews on a given topic. Example topics are “performance of Toyota Ca
mry” and “sound quality of ipod nano”.')
descrip.insert(358,'This data set contains user reviews of cars and and
 hotels collected from Tripadvisor (~259,000 reviews) and Edmunds (~42,
230 reviews).')
descrip.insert(364,'The goal is to model mutant p53 transcriptional act
ivity (active vs inactive) based on data extracted from biophysical sim
ulations.')
descrip.insert(369,'Data collected from car parks in Birmingham that ar
e operated by NCP from Birmingham City Council. UK Open Government Lice
nce (OGL). https://data.birmingham.gov.uk/dataset/birmingham-parking')
```

```
descrip.insert(393,'PPG-DaLiA contains data from 15 subjects wearing ph
ysiological and motion sensors, providing a PPG dataset for motion comp
ensation and heart rate estimation in Daily Life Activities.')
descrip.insert(413,'The “real estate valuation” is a regression proble
m. The market historical data set of real estate valuation are collecte
d from Sindian Dist., New Taipei City, Taiwan.')
descrip.insert(423,'This dataset is created by reading out 200 files fr
om the 10 largest Reuters classes and using an Automatic Speech Recogni
tion system to create corresponding transcriptions.')
descrip.insert(436,'The data describe the problem of high energy (highe
r than 10^4 J) seismic bumps forecasting in a coal mine. Data come from
 two of longwalls located in a Polish coal mine.')
descrip.insert(437,'The SELFBACK dataset is a Human Activity Recognitio
n Dataset of 9 activity classes recorded with two tri-axial acceleromet
ers.')
descrip.insert(455,'Activity recognition data set built from the record
ings of 30 subjects performing basic activities and postural transition
s while carrying a waist-mounted smartphone with embedded inertial sens
ors. ')
descrip.insert(479,'A dataset of steel plates’ faults, classified into
 7 different types. The goal was to train machine learning for automati
c pattern recognition.')
descrip.insert(480,'The dataset is provided by the “Trialto Latvia LT
D”, the third-party logistics operator. Each observation stands for a d
istinct type of item for sale.')
descrip.insert(482,'Stone flakes are waste products of the stone tool p
roduction in the prehistoric era. The variables are means of geometric
 and stylistic features of the flakes contained in different inventorie
s.')
descrip.insert(494,'An accurate dataset describing trajectories perform
ed by all the 442 taxis running in the city of Porto, in Portugal.')
descrip.insert(509,'UbiqLog is the smartphone lifelogging tool that run
s on the smartphone of 35 users for about 2 months.')
descrip.insert(524,' The dataset collects data from an Android smartpho
ne positioned in the chest pocket from 22 participants walking in the w
ild over a predefined path.')
descrip.insert(540,'')
descrip.insert(547,'Two datasets are included, related to red and white
 vinho verde wine samples, from the north of Portugal. The goal is to m
```

```
odel wine quality based on physicochemical tests (see [Cortez et al., 2
009], http://www3.dsi.uminho.pt/pcortez/wine/).')
descrip.insert(554,' The datasets are taken from top 2 Indian cooking c
hannel named Nisha Madhulika channel and Kabita's Kitchen channel. The
 data set is in Hinglish Language.')
# no supe como mas hacerlo :<
```

In [141]:
```
len(descrip)
```

Out[141]:
```
559
```

In [142]:
```
bb= re.findall('''">([\w''  "0-9/?^.@+\,;*$:!@()_+—&-<>/]*)</a></b>''',
 str(description))
bb
```

Out[142]:
```
['2.4 GHZ Indoor Channel Measurements',
 '3D Road Network (North Jutland, Denmark)',
 '3W dataset',
 ': Simulated Data set of Iraqi tourism places',
 'A study of  Asian Religious and Biblical Texts',
 'AAAI 2013 Accepted Papers',
 'AAAI 2014 Accepted Papers',
 'Abalone',
 'Abscisic Acid Signaling Network',
 'Absenteeism at work',
 'Activities of Daily Living (ADLs) Recognition Using Binary Sensors',
 'Activity Recognition from Single Chest-Mounted Accelerometer',
 'Activity Recognition system based on Multisensor data fusion (AReM)',
 'Activity recognition using wearable physiological measurements',
 'Activity recognition with healthy older people using a batteryless we
arable sensor',
 'Acute Inflammations',
 'Adult',
 'Air Quality',
 'Air quality',
 'Airfoil Self-Noise',
 'Alcohol QCM Sensor Dataset',
 'Algerian Forest Fires Dataset  ',
 'Amazon Access Samples',
```

```
'Amazon Commerce reviews set',
'Amphibians',
'Annealing',
'Anonymous Microsoft Web Data',
'Anuran Calls (MFCCs)',
'Apartment for rent classified',
'Appliances energy prediction',
'APS Failure at Scania Trucks',
'Arcene',
'Arrhythmia',
'Artificial Characters',
'Audiology (Original)',
'Audiology (Standardized)',
'Audit Data',
'Australian Sign Language signs',
'Australian Sign Language signs (High Quality)',
'Autism Screening Adult',
'Autistic Spectrum Disorder Screening Data for Adolescent   ',
'Autistic Spectrum Disorder Screening Data for Children  ',
'Auto MPG',
'Automobile',
'AutoUniv',
'Avila',
'Bach Choral Harmony',
'Bach Chorales',
'Badges',
'Bag of Words',
'Balance Scale',
'Balloons',
'Bank Marketing',
'banknote authentication',
'Bar Crawl: Detecting Heavy Drinking',
'Bar Crawl: Detecting Heavy Drinking',
'BAUM-1',
'BAUM-2',
'Behavior of the urban traffic of the city of Sao Paulo in Brazil',
'Beijing Multi-Site Air-Quality Data',
'Beijing PM2.5 Data',
'Bias correction of numerical prediction model temperature forecast',
```

```
 'Bike Sharing Dataset',
 'BitcoinHeistRansomwareAddressDataset',
 'BLE RSSI Dataset for Indoor localization and Navigation',
 'BlogFeedback',
 'BLOGGER',
 'Blood Transfusion Service Center',
 'Bone marrow transplant: children',
 'Breast Cancer',
 'Breast Cancer Coimbra',
 'Breast Cancer Wisconsin (Diagnostic)',
 'Breast Cancer Wisconsin (Original)',
 'Breast Cancer Wisconsin (Prognostic)',
 'Breast Tissue',
 'Breath Metabolomics',
 'BuddyMove Data Set',
 'Burst Header Packet (BHP) flooding attack on Optical Burst Switching
(OBS) Network',
 'Buzz in social media ',
 'Caesarian Section Classification Dataset',
 'CalIt2 Building People Counts',
 'Car Evaluation',
 'Carbon Nanotubes',
 'Cardiotocography',
 'Cargo 2000 Freight Tracking and Tracing',
 'Census Income',
 'Census-Income (KDD)',
 'Cervical cancer (Risk Factors)',
 'Cervical Cancer Behavior Risk',
 'Challenger USA Space Shuttle O-Ring',
 'Character Font Images',
 'Character Trajectories',
 'Chess (Domain Theories)',
 'Chess (King-Rook vs. King)',
 'Chess (King-Rook vs. King-Knight)',
 'Chess (King-Rook vs. King-Pawn)',
 'chestnut – LARVIC',
 'chipseq',
 'Chronic_Kidney_Disease',
 'clickstream data for online shopping',
```

```
 'Climate Model Simulation Crashes',
 'CLINC150',
 'Cloud',
 'CMU Face Images',
 'CNAE-9',
 'CNNpred: CNN-based stock market prediction using a diverse set of var
iables',
 'Codon usage',
 'Coil 1999 Competition Data',
 'Combined Cycle Power Plant',
 'Communities and Crime',
 'Communities and Crime Unnormalized',
 'Computer Hardware',
 'Concrete Compressive Strength',
 'Concrete Slump Test',
 'Condition Based Maintenance of Naval Propulsion Plants',
 'Condition monitoring of hydraulic systems',
 'Congressional Voting Records',
 'Connect-4',
 'Connectionist Bench (Nettalk Corpus)',
 'Connectionist Bench (Sonar, Mines vs. Rocks)',
 'Connectionist Bench (Vowel Recognition - Deterding Data)',
 'Container Crane Controller Data Set',
 'Contraceptive Method Choice',
 'Corel Image Features',
 'Covertype',
 'COVID-19 Surveillance',
 'Credit Approval',
 'Crop mapping using fused optical-radar data set',
 'Crowdsourced Mapping',
 'Cryotherapy Dataset ',
 'CSM (Conventional and Social Media Movies) Dataset 2014 and 2015',
 'Cuff-Less Blood Pressure Estimation',
 'Cylinder Bands',
 'Daily and Sports Activities',
 'Daily Demand Forecasting Orders',
 'Daphnet Freezing of Gait',
 'Data for Software Engineering Teamwork Assessment in Education Settin
g',
```

```
 'Dataset for ADL Recognition with Wrist-worn Accelerometer',
 'Dataset for Sensorless Drive Diagnosis',
 'DBWorld e-mails',
 'Deepfakes: Medical Image Tamper Detection',
 'default of credit card clients',
 'DeliciousMIL: A Data Set for Multi-Label Multi-Instance Learning with
Instance Labels',
 'Demand Forecasting for a store',
 'Demospongiae',
 'Dermatology',
 'Detect Malacious Executable(AntiVirus)',
 'Detect Malware Types',
 'detection_of_IoT_botnet_attacks_N_BaIoT',
 'Devanagari Handwritten Character Dataset',
 'Dexter',
 'DGP2 - The Second Data Generation Program',
 'Diabetes',
 'Diabetes 130-US hospitals for years 1999-2008',
 'Diabetic Retinopathy Debrecen Data Set',
 'Discrete Tone Image Dataset',
 'Dishonest Internet users Dataset',
 'Divorce Predictors data set',
 'Divorce Predictors data set',
 'Document Understanding',
 'Dodgers Loop Sensor',
 'Dorothea',
 'Dota2 Games Results',
 'Dow Jones Index',
 'Dresses_Attribute_Sales',
 'DrivFace',
 'Drug consumption (quantified)',
 'Drug Review Dataset (Druglib.com)',
 'Drug Review Dataset (Drugs.com)',
 'DSRC Vehicle Communications',
 'Dynamic Features of VirusShare Executables',
 'E. Coli Genes',
 'Early biomarkers of Parkinson’s disease based on natural connected sp
eech Data Set ',
 'Early biomarkers of Parkinson's disease based on natural connected sp
```

```
eech',
 'Early stage diabetes risk prediction dataset.',
 'EBL Domain Theories',
 'Echocardiogram',
 'Eco-hotel',
 'Ecoli',
 'Economic Sanctions',
 'Educational Process Mining (EPM): A Learning Analytics Data Set',
 'EEG Database',
 'EEG Eye State',
 'EEG Steady-State Visual Evoked Potential Signals',
 'El Nino',
 'Electrical Grid Stability Simulated Data ',
 'ElectricityLoadDiagrams20112014',
 'EMG data for gestures',
 'EMG dataset in Lower Limb',
 'EMG Physical Action Data Set',
 'Energy efficiency',
 'Entree Chicago Recommendation Data',
 'Epileptic Seizure Recognition',
 'Estimation of obesity levels based on eating habits and physical cond
ition ',
 'Exasens',
 'Exasens',
 'extention of Z-Alizadeh sani dataset',
 'Facebook Comment Volume Dataset',
 'Facebook Large Page-Page Network',
 'Facebook Live Sellers in Thailand',
 'Facebook metrics',
 'Farm Ads',
 'Fertility',
 'Firm-Teacher_Clave-Direction_Classification',
 'First-order theorem proving',
 'Flags',
 'FMA: A Dataset For Music Analysis',
 'Folio',
 'Forest Fires',
 'Forest type mapping',
 'Function Finding',
```

```
 'Gas Sensor Array Drift Dataset',
 'Gas Sensor Array Drift Dataset at Different Concentrations',
 'Gas sensor array exposed to turbulent gas mixtures',
 'Gas sensor array temperature modulation',
 'Gas sensor array under dynamic gas mixtures',
 'Gas sensor array under flow modulation',
 'Gas sensor arrays in open sampling settings',
 'Gas sensors for home activity monitoring',
 'Gas Turbine CO and NOx Emission Data Set',
 'Gastrointestinal Lesions in Regular Colonoscopy',
 'gene expression cancer RNA-Seq',
 'Geo-Magnetic field and WLAN dataset for indoor localisation from wris
tband and smartphone',
 'Geographical Original of Music',
 'Gesture Phase Segmentation',
 'Gisette',
 'Glass Identification',
 'GNFUV Unmanned Surface Vehicles Sensor Data',
 'GNFUV Unmanned Surface Vehicles Sensor Data Set 2',
 'GPS Trajectories',
 'Grammatical Facial Expressions',
 'Greenhouse Gas Observing Network',
 'Guitar Chords finger positions',
 "Haberman's Survival",
 'Hayes-Roth',
 'HCC Survival',
 'HCV data',
 'Health News in Twitter',
 'Heart Disease',
 'Heart failure clinical records',
 'Hepatitis',
 'Hepatitis C Virus (HCV) for Egyptian patients',
 'HEPMASS',
 'Heterogeneity Activity Recognition',
 'HIGGS',
 'Hill-Valley',
 'HIV-1 protease cleavage',
 'Horse Colic',
 'Horton General Hospital',
```

```
 'HTRU2',
 'Human Activity Recognition from Continuous Ambient Sensor Data',
 'Human Activity Recognition Using Smartphones',
 'Hybrid Indoor Positioning Dataset from WiFi RSSI, Bluetooth and magne
tometer',
 'ICMLA 2014 Accepted Papers Data Set',
 'ICU',
 'IDA2016Challenge',
 'IIWA14-R820-Gazebo-Dataset-10Trajectories',
 'ILPD (Indian Liver Patient Dataset)',
 'Image Segmentation',
 'Immunotherapy Dataset',
 'Improved Spiral Test Using Digitized Graphics Tablet for Monitoring P
arkinson's Disease',
 'Incident management process enriched event log',
 'Individual household electric power consumption',
 'Indoor User Movement Prediction from RSS data',
 'Insurance Company Benchmark (COIL 2000)',
 'Intelligent Media Accelerometer and Gyroscope (IM-AccGyro) Dataset',
 'Internet Advertisements',
 'Internet Firewall Data',
 'Internet Usage Data',
 'Ionosphere',
 'IPUMS Census Database',
 'Iranian Churn Dataset',
 'Iris',
 'ISOLET',
 'ISTANBUL STOCK EXCHANGE',
 'Japanese Credit Screening',
 'Japanese Vowels',
 'KASANDR',
 'KDC-4007 dataset Collection',
 'KDD Cup 1998 Data',
 'KDD Cup 1999 Data',
 'KEGG Metabolic Reaction Network (Undirected)',
 'KEGG Metabolic Relation Network (Directed)',
 'Kinship',
 'Kitsune Network Attack Dataset',
 'Labor Relations',
```

```
'Las Vegas Strip',
'Leaf',
'LED Display Domain',
'Legal Case Reports',
'Lenses',
'Letter Recognition',
'Libras Movement',
'Liver Disorders',
'Localization Data for Person Activity',
'Logic Theorist',
'Low Resolution Spectrometer',
'LSVT Voice Rehabilitation',
'Lung Cancer',
'Lymphography',
'M. Tuberculosis Genes',
'Machine Learning based ZZAlpha Ltd. Stock Recommendations 2012-2014',
'Madelon',
'MAGIC Gamma Telescope',
'Malware static and dynamic features VxHeaven and Virus Total',
'Mammographic Mass',
'Mechanical Analysis',
'Mesothelioma’s disease data set ',
'Meta-data',
'Metro Interstate Traffic Volume',
'MEU-Mobile KSD',
'MEx',
'MHEALTH Dataset',
'Mice Protein Expression',
'microblogPCU',
'MicroMass',
'MiniBooNE particle identification',
'Miskolc IIS Hybrid IPS',
'Mobile Robots',
'MoCap Hand Postures',
'Molecular Biology (Promoter Gene Sequences)',
'Molecular Biology (Protein Secondary Structure)',
'Molecular Biology (Splice-junction Gene Sequences)',
"MONK's Problems",
'Monolithic Columns in Troad and Mysia Region',
```

```
'Moral Reasoner',
'Motion Capture Hand Postures',
'Movie',
'MSNBC.com Anonymous Web Data',
'Mturk User-Perceived Clusters over Images',
'Multimodal Damage Identification for Humanitarian Computing',
'Multiple Features',
'Mushroom',
'Musk (Version 1)',
'Musk (Version 2)',
'Nasarian CAD Dataset',
'News Aggregator',
'News Popularity in Multiple Social Media Platforms',
'Newspaper and magazine images segmentation dataset',
'NIPS Conference Papers 1987-2015',
'NoisyOffice',
'Nomao',
'Northix',
'NSF Research Award Abstracts 1990-2003',
'Nursery',
'NYSK',
'Occupancy Detection ',
'OCT data &amp;amp; Color Fundus Images of Left &amp;amp; Right Eyes',
'One-hundred plant species leaves data set',
'Online Handwritten Assamese Characters Dataset',
'Online News Popularity',
'Online Retail',
'Online Retail II',
'Online Shoppers Purchasing Intention Dataset',
'Online Video Characteristics and Transcoding Time Dataset',
'Open University Learning Analytics dataset',
'Opinion Corpus for Lebanese Arabic Reviews (OCLAR)',
'Opinosis Opinion ⁄ Review',
'OpinRank Review Dataset',
'OPPORTUNITY Activity Recognition',
'Optical Interconnection Network ',
'Optical Recognition of Handwritten Digits',
'Othello Domain Theory',
'Ozone Level Detection',
```

```
'p53 Mutants',
'Page Blocks Classification',
'PAMAP2 Physical Activity Monitoring',
'PANDOR',
'Paper Reviews',
'Parking Birmingham',
'Parkinson Dataset with replicated acoustic features ',
'Parkinson Disease Spiral Drawings Using Digitized Graphics Tablet',
'Parkinson Speech Dataset with  Multiple Types of Sound Recordings',
"Parkinson's Disease Classification",
'Parkinsons',
'Parkinsons Telemonitoring',
'Pedestrian in Traffic Dataset',
'PEMS-SF',
'Pen-Based Recognition of Handwritten Digits',
'Perfume Data',
'Person Classification Gait Data',
'Phishing Websites',
'Physical Unclonable Functions',
'Physicochemical Properties of Protein Tertiary Structure',
'Pioneer-1 Mobile Robot Data',
'Pittsburgh Bridges',
'Planning Relax',
'Plants',
'PM2.5 Data of Five Chinese Cities',
'PMU-UD',
'Poker Hand',
'Polish companies bankruptcy data',
'Post-Operative Patient',
'PPG-DaLiA',
'Predict keywords activities in a online social media',
'Primary Tumor',
'Prodigy',
'Protein Data',
'Pseudo Periodic Synthetic Time Series',
'PubChem Bioassay Data',
'QSAR androgen receptor',
'QSAR aquatic toxicity',
'QSAR Bioconcentration classes dataset',
```

```
 'QSAR biodegradation',
 'QSAR fish bioconcentration factor (BCF)',
 'QSAR fish toxicity',
 'QSAR oral toxicity',
 'QtyT40I10D100K',
 'Quadruped Mammals',
 'Qualitative Structure Activity Relationships',
 'Qualitative_Bankruptcy',
 'Quality Assessment of Digital Colposcopies',
 'Query Analytics Workloads Dataset',
 'Real estate valuation data set',
 'Real-time Election Results: Portugal 2019',
 'REALDISP Activity Recognition Dataset',
 'Record Linkage Comparison Patterns',
 'Refractive errors',
 'Relative location of CT slices on axial axis',
 'Repeat Consumption Matrices',
 'Residential Building Data Set',
 'Restaurant &amp;amp; consumer data',
 'Reuters RCV1 RCV2 Multilingual, Multiview Text Categorization Test co
llection',
 'Reuters Transcribed Subset',
 'Reuters-21578 Text Categorization Collection',
 'Reuter_50_50',
 'Rice (Cammeo and Osmancik)',
 'Rice Leaf Diseases',
 'Robot Execution Failures',
 'Roman Urdu Data Set',
 'Russian Corpus of Biographical Texts',
 'Sales_Transactions_Dataset_Weekly',
 'Sattriya_Dance_Single_Hand_Gestures Dataset',
 'SCADI',
 'SECOM',
 'seeds',
 'seismic-bumps',
 'selfBACK',
 'Semeion Handwritten Digit',
 'sEMG for Basic Hand movements',
 'Sentence Classification',
```

```
 'Sentiment Labelled Sentences',
 'Seoul Bike Sharing Demand',
 "ser Knowledge Modeling Data (Students' Knowledge Levels on DC Electri
cal Machines)",
 'Servo',
 'SGEMM GPU kernel performance',
 'Shill Bidding Dataset',
 'Shoulder Implant X-Ray Manufacturer Classification',
 'Shoulder Implant X-Ray Manufacturer Classification',
 'Shuttle Landing Control',
 'SIFT10M',
 'Simulated Falls and Daily Living Activities Data Set',
 'SkillCraft1 Master Table Dataset',
 'Skin Segmentation',
 'Smartphone Dataset for Human Activity Recognition (HAR) in Ambient As
sisted Living (AAL)',
 'Smartphone-Based Recognition of Human Activities and Postural Transit
ions',
 'SML2010',
 'SMS Spam Collection',
 'Solar Flare',
 'Somerville Happiness Survey',
 'South German Credit',
 'South German Credit (UPDATE)',
 'Soybean (Large)',
 'Soybean (Small)',
 'Spambase',
 'Speaker Accent Recognition',
 'SPECT Heart',
 'SPECTF Heart',
 'Spoken Arabic Digit',
 'Sponge',
 'Sports articles for objectivity analysis',
 'Statlog (Australian Credit Approval)',
 'Statlog (German Credit Data)',
 'Statlog (Heart)',
 'Statlog (Image Segmentation)',
 'Statlog (Landsat Satellite)',
 'Statlog (Shuttle)',
```

```
'Statlog (Vehicle Silhouettes)',
'Statlog Project',
'Steel Plates Faults',
'Stock keeping units',
'Stock portfolio performance',
'StoneFlakes',
'Student Academics Performance',
'Student Loan Relational',
'Student Performance',
'Superconductivty Data',
'SUSY',
'Swarm Behaviour',
'Synthetic Control Chart Time Series',
'Syskill and Webert Web Page Ratings',
'Taiwanese Bankruptcy Prediction',
'Tamilnadu Electricity Board Hourly Readings',
'Tarvel Review Ratings',
'Taxi Service Trajectory - Prediction Challenge, ECML PKDD 2015',
'Teaching Assistant Evaluation',
'Tennis Major Tournament Match Statistics',
'Thoracic Surgery Data',
'Thyroid Disease',
'Tic-Tac-Toe Endgame',
'Trains',
'Travel Reviews',
'TTC-3600: Benchmark dataset for Turkish text categorization',
'Turkish Spam V01',
'Turkiye Student Evaluation',
'TV News Channel Commercial Detection Dataset',
'Twenty Newsgroups',
'Twin gas sensor arrays',
'Twitter Data set for Arabic Sentiment Analysis',
'UbiqLog (smartphone lifelogging)',
'UJI Pen Characters',
'UJI Pen Characters (Version 2)',
'UJIIndoorLoc',
'UJIIndoorLoc-Mag',
'Ultrasonic flowmeter diagnostics',
'Undocumented',
```

```
 'University',
 'University of Tehran Question Dataset 2016 (UTQD.2016)',
 'UNIX User Data',
 'Unmanned Aerial Vehicle (UAV) Intrusion Detection',
 'Urban Land Cover',
 'UrbanGB, urban road accidents coordinates labelled by the urban cente
r',
 'URL Reputation',
 'US Census Data (1990)',
 'User Identification From Walking Activity',
 'User Knowledge Modeling',
 'User Profiling and Abusive Language Detection Dataset',
 'USPTO Algorithm Challenge, run by NASA-Harvard Tournament Lab and Top
Coder    Problem: Pat',
 'Vehicle routing and scheduling problems',
 'Vertebral Column',
 'Vicon Physical Action Data Set',
 'Victorian Era Authorship Attribution',
 'Volcanoes on Venus - JARtool experiment',
 'Wall-Following Robot Navigation Data',
 'Water Treatment Plant',
 'Wave Energy Converters',
 'Wave Energy Converters',
 'Waveform Database Generator (Version 1)',
 'Waveform Database Generator (Version 2)',
 'Wearable Computing: Classification of Body Postures and Movements (PU
C-Rio)',
 'Website Phishing',
 'Weight Lifting Exercises monitored with Inertial Measurement Units',
 'WESAD (Wearable Stress and Affect Detection)',
 'Wholesale customers',
 'wiki4HE',
 'Wilt',
 'Wine',
 'Wine Quality',
 'Wireless Indoor Localization',
 'WISDM Smartphone and Smartwatch Activity and Biometrics Dataset ',
 'Yacht Hydrodynamics',
 'YearPredictionMSD',
```

```
 'Yeast',
 'YouTube Comedy Slam Preference Data',
 'Youtube cookery channels viewers comments in Hinglish',
 'YouTube Multiview Video Games Dataset',
 'YouTube Spam Collection',
 'Z-Alizadeh Sani',
 'Zoo']
```

In [143]:
```
len(bb)
```

Out[143]: 559

In [144]:
```
import pandas as pd
dframe = {'Description':descrip}
ohalfdf = pd.DataFrame(dframe,index =bb)
ohalfdf
```

Out[144]:

	Description
2.4 GHZ Indoor Channel Measurements	Measurement of the S21,consists of 10 sweeps, ...
3D Road Network (North Jutland, Denmark)	3D road network with highly accurate elevation...
3W dataset	The first realistic and public dataset with ra...
: Simulated Data set of Iraqi tourism places	Simulated Data set of Iraqi tourism places wit...
A study of Asian Religious and Biblical Texts	Mainly from Project Gutenberg, we combine Upan...
...	...
Youtube cookery channels viewers comments in Hinglish	The datasets are taken from top 2 Indian cook...
YouTube Multiview Video Games Dataset	This dataset contains about 120k instances, ea...
YouTube Spam Collection	It is a public set of comments collected for s...
Z-Alizadeh Sani	It was collected for CAD diagnosis.
Zoo	Artificial, 7 classes of animals

559 rows × 1 columns

In [152]:
```
final_data= pd.merge(halfdf,ohalfdf,how='inner' ,left_index=True, right
_index=True)
final_data
```

Out[152]:

	Data Types	Default Task	Attribute Types	Num Instances	Num Attributes	Year	
2.4 GHZ Indoor Channel Measurements	Multivariate	Classification	Real	7840	5	2018	https://archive.ics.u
3D Road Network (North Jutland, Denmark)	Sequential, Text	Regression, Clustering	Real	434874	4	2013	https://archive.ics.u
3W dataset	Multivariate, Time-Series	Classification, Clustering	Integer, Real	1984	8	2019	https://archive.ics.u
: Simulated Data set of Iraqi tourism places	Multivariate	Classification, Clustering		232	16	2020	https://archive.ics.u
A study of Asian Religious and Biblical Texts	Multivariate, Text	Classification, Clustering	Integer	590	8265	2019	https://archive.ics.
...	...	...	...	...	...	...	
seeds	Multivariate	Classification, Clustering	Real	210	7	2012	https://archive.i

	Data Types	Default Task	Attribute Types	Num Instances	Num Attributes	Year	
seismic-bumps	Multivariate	Classification	Real	2584	19	2013	https://archive.ics
selfBACK	Time-Series	Classification, Clustering	Real	26136	6	2020	https://archive.ics.u
ser Knowledge Modeling Data (Students' Knowledge Levels on DC Electrical Machines)	Multivariate	Classification	Real	403	5	2013	https://archive.ics.u
wiki4HE	Multivariate	Regression, Clustering, Causal-Discovery		913	53	2015	https://archive.ics

569 rows × 8 columns

salen 10 filas más, de seguro son bases con misspelling (porque halfdf y ohalfdf tienen 559 filas)