

DTS 201 – Introduction to Data Science

Week 2 Practical (PH): Data Collection, Cleaning, and Quality Assessment

Instructor: Dr. Sakinat Folorunso

Title: Associate Professor of AI Systems and FAIR Data

Department: Computer Sciences, Olabisi Onabanjo University, Ago-Iwoye, Ogun State, Nigeria

Practical Learning Objectives (Week 2)

By the end of this practical session, you should be able to:

- Collect data from multiple sources (manual entry, web files, shared documents).
- Import and combine raw datasets using Python (CSV, Excel, TXT).
- Detect common data quality problems (missing values, inconsistent labels, wrong data types).
- Clean and standardize real-world data for analysis.
- Apply basic validation checks to ensure the correctness and reliability of your dataset.

 **This notebook is highly student-centred.** You will perform each task hands-on, modify the code, interpret results, and submit your own reflections.

1. Getting Started: Where Will Your Data Come From?

In this practical, you will practice collecting datasets from these common sources:

1. **Files downloaded from the web** (CSV, XLSX).
2. **Manual entry**, where you create a DataFrame from scratch.
3. **Shared departmental data** (provided by your lecturer).

You may work with:

- Your own dataset
- The sample datasets provided
- Or create a small dataset of your own

1.1 Optional: Mount Google Drive (If Using Colab)

```
In [ ]: # Run this only if using Google Colab AND your dataset is stored on Google Drive.

# Import tool for mounting Google Drive into Colab
from google.colab import drive

# Mount your Google Drive - follow the instructions when you run this
drive.mount('/content/drive')

# After mounting, your files become accessible under '/content/drive/MyDrive/'
```

Mounted at /content/drive

2. Importing Required Libraries

```
In [4]: # pandas: primary library for working with data tables
import pandas as pd

# Import the pandas Library for working with tabular data (rows and columns)
import pandas as pd

# Import numpy for numerical operations (we may need it later)
import numpy as np

# Import matplotlib for basic plotting (optional for this week but useful generally)
import matplotlib.pyplot as plt

# Configure matplotlib so that any plots appear inside the notebook
%matplotlib inline

import seaborn as sns
```

3. Collecting and Importing Raw Datasets

We will work with datasets that may contain:

- Typos
- Missing values
- Inconsistent labels
- Duplicate entries
- Wrong data types

You will load three datasets:

- A **CSV** file
- An **Excel** file
- A **TXT** / delimited file

3.1 Importing a CSV File

```
In [5]: # Replace 'raw_students.csv' with your real CSV file name
csv_path = '../data/Sales_data.csv' # Updated path for Local setup

# Load CSV file
df_csv = pd.read_csv(csv_path) # Reads CSV into DataFrame

# View sample rows
df_csv.head() # Look at the first few rows to inspect raw structure
```

Out[5]:

	product_id	product_name	category	discounted_
0	B07JW9H4J1	Wayona Nylon Braided USB to Lightning Fast Charging Cha...	Computers&Accessories Accessories&Peripherals ...	
1	B098NS6PVG	Ambrane Unbreakable 60W / 3A Fast Charging 1.5...	Computers&Accessories Accessories&Peripherals ...	
2	B096MSW6CT	Sounce Fast Phone Charging Cable & Data Sync U...	Computers&Accessories Accessories&Peripherals ...	
3	B08HDJ86NZ	boAt Deuce USB 300 2 in 1 Type-C & Micro USB S...	Computers&Accessories Accessories&Peripherals ...	
4	B08CF3B7N1	Portronics Konnect L 1.2M Fast Charging 3A 8 P...	Computers&Accessories Accessories&Peripherals ...	

3.2 Importing an Excel File

```
In [6]: # Replace 'raw_sales.xlsx' with your Excel file
excel_path = '../data/Sales_data.csv' # Using CSV instead of Excel

# Read CSV file - same data, different format
df_excel = pd.read_csv(excel_path)

df_excel.head()
```

Out[6]:

	product_id	product_name	category	discounted_
0	B07JW9H4J1	Wayona Nylon Braided USB to Lightning Fast Charger	Computers&Accessories Accessories&Peripherals ...	
1	B098NS6PVG	Ambrane Unbreakable 60W / 3A Fast Charging	Computers&Accessories Accessories&Peripherals ...	
2	B096MSW6CT	Source Fast Phone Charging Cable & Data Sync U...	Computers&Accessories Accessories&Peripherals ...	
3	B08HDJ86NZ	boAt Deuce USB 300 2 in 1 Type-C & Micro USB S...	Computers&Accessories Accessories&Peripherals ...	
4	B08CF3B7N1	Portronics Konnect L 1.2M Fast Charging 3A 8 P...	Computers&Accessories Accessories&Peripherals ...	

◀ ▶

3.3 Importing a TXT / Delimited File

In [7]:

```
txt_path = '../data/Sales_data.csv' # Using same file for TXT example

# Load a text file separated by commas or tabs
df_txt = pd.read_csv(txt_path, delimiter=",") # Using comma delimiter

df_txt.head()
```

Out[7]:

	product_id	product_name	category	discounted_
0	B07JW9H4J1	Wayona Nylon Braided USB to Lightning Fast Charger	Computers&Accessories Accessories&Peripherals ...	
1	B098NS6PVG	Ambrane Unbreakable 60W / 3A Fast Charging	Computers&Accessories Accessories&Peripherals ...	
2	B096MSW6CT	Source Fast Phone Charging Cable & Data Sync U...	Computers&Accessories Accessories&Peripherals ...	
3	B08HDJ86NZ	boAt Deuce USB 300 2 in 1 Type-C & Micro USB S...	Computers&Accessories Accessories&Peripherals ...	
4	B08CF3B7N1	Portronics Konnect L 1.2M Fast Charging 3A 8 P...	Computers&Accessories Accessories&Peripherals ...	

In [6]:

```
txt_path = '../data/Sales_data.csv'

# Read and display the first 10 lines with proper encoding
with open(txt_path, 'r', encoding='utf-8') as f:
    for i, line in enumerate(f):
        print(f"Line {i+1}: {line.strip()}")
        if i == 9: # Display only the first 10 lines
            break
```

Line 1: product_id,product_name,category,discounted_price,actual_price,discount_per centage,rating,rating_count,about_product,user_id,user_name,review_id,review_title,review_content,img_link,product_link

Line 2: B07JW9H4J1,"Wayona Nylon Braided USB to Lightning Fast Charging and Data Sync Cable Compatible for iPhone 13, 12,11, X, 8, 7, 6, 5, iPad Air, Pro, Mini (3 FT Pack of 1, Grey)",Computers&Accessories|Accessories&Peripherals|Cables&Accessories|Cables|USBCables,₹399,"₹1,099",64%,4.2,"24,269","High Compatibility : Compatible With iPhone 12, 11, X/XsMax/Xr ,iPhone 8/8 Plus,iPhone 7/7 Plus,iPhone 6s/6s Plus,iPhone 6/6 Plus,iPhone 5/5s/5c/se,iPad Pro,iPad Air 1/2,iPad mini 1/2/3,iPod nano7,iPod touch and more apple devices.|Fast Charge&Data Sync : It can charge and sync simultaneously at a rapid speed, Compatible with any charging adaptor, multi-port charging station or power bank.|Durability : Durable nylon braided design with premium aluminum housing and toughened nylon fiber wound tightly around the cord lending it superior durability and adding a bit to its flexibility.|High Security Level : It is designed to fully protect your device from damaging excessive current.Copper core thick+Multi layer shielding, Anti-interference, Protective circuit equipment.|WARRANTY: 12 months warranty and friendly customer services, ensures the long-time enjoyment of your purchase. If you meet any question or problem, please don't hesitate to contact us.", "AG3D604STAQKAY2UVGEUV46KN35Q,AHMY5CWJMMK5BJRBSNLYT3ONILA,AHCTC6ULH4XB6YHDY6PCH2R772LQ,AGYHHIERNXKA6P5T7CZLXKVPT7IQ,AG40GOFWXJZTQ2HKYIOCOY3KXF2Q,AENGU523SXMO57JPDTW52PNNVWGQ,AEQJHCVTNINBS4FKTBGQRQTGTE5Q,AFC3FFC5PKFF5PMA52S3VCHOZ5FQ", "Manav,Adarsh gupta,Sundeep,S.Sayeed Ahmed,jaspreet singh,Khaja moin,Anand,S.ARUMUGAM", "R3HXWT0LRP0NMF,R2AJM3LFTLZHFO,R6AQJGUP6P86,R1KD19HEDV0OR,R3C02RMYQMK6FC,R39GQRVBUZBWGY,R2K9ED0E15QIRJ,R3O17YT648TL8I", "Satisfied,Charging is really fast,Value for money,Product review,Good quality,Good product,Good Product,As of now seems good", "Looks durable Charging is fine tooNo complains,Charging is really fast, good product.,Till now satisfied with the quality.,This is a good product . The charging speed is slower than the original iPhone cable,Good quality, would recommend,https://m.media-amazon.com/images/W/WEBP_402378-T1/images/I/81---F1ZgHL._SY88.jpg,Product had worked well till date and was having no issue.Cable is also sturdy enough...Have asked for replacement and company is doing the same...,Value for money", https://m.media-amazon.com/images/W/WEBP_402378-T1/images/I/51UsScvHQNL._SX300_SY300_QL70_FMwebp_.jpg,https://www.amazon.in/Wayona-Braided-WN3LG1-Syncing-Charging/dp/B07JW9H4J1/ref=sr_1_1?qid=1672909124&s=electronics&sr=1-1

Line 3: B098NS6PVG,"Ambrane Unbreakable 60W / 3A Fast Charging 1.5m Braided Type C Cable for Smartphones, Tablets, Laptops & other Type C devices, PD Technology, 480Mbps Data Sync, Quick Charge 3.0 (RCT15A, Black)",Computers&Accessories|Accessories&Peripherals|Cables&Accessories|Cables|USBCables,₹199,₹349,43%,4.0,"43,994", "Compatible with all Type C enabled devices, be it an android smartphone (Mi, Samsung, Oppo, Vivo, Realme, OnePlus, etc), tablet, laptop (Macbook, Chromebook, etc)|Supports Quick Charging (2.0/3.0)|Unbreakable - Made of special braided outer with rugged interior bindings, it is ultra-durable cable that won't be affected by daily rough usage|Ideal Length - It has ideal length of 1.5 meters which is neither too short like your typical 1meter cable or too long like a 2meters cable|Supports maximum 3A fast charging and 480 Mbps data transfer speed|6 months manufacturer warranty from the date of purchase", "AECPFYFQVRUWC3KGNLJ1OREFP5LQ,AGYYVPDD7YG7FYNBXNGXZJT525AQ,AHONIZU3CIEHQIGQ6R2VFRSBXOO,AFPHD2CRPDZMWMBL7WXRSVYWS5JA,AEZ346GX3HJ404XNRPHCNHXQURMQ,AEPSWFNECKO34PUC7I56ITGXR6Q,AHWVEHR5DYLVFT02KF3IZATFQSWQ,AH4QT33M55677I7ISQOAKEQWACYQ", "ArdKn,Nirbhay kumar,Sagar Viswanathan,Asp,Placeholder,BharanI,sonia,Niam", "RGIQEG07R9HS2,R1SMWZQ86XIN8U,R2J3Y1WL29GWDE,RYGGS0M09S3KY,R17KQRUTAN5DKS,R3AAQGS6HP2QUK,R1HDNOG6T02CCA,R3PHKXYA5AFE0U", "A Good Braided Cable for Your Type C Device,Good quality product from ambrane,Super cable,As,Good quality,Good product,its good,Good quality for the price but one issue with my unit","I ordered this cable to connect my phone to Android Auto of car. The cable is really strong and the connection ports are really well made. I already has a Micro USB cable from Ambrane and it's still in good shape. I connected my phone to the car using the cable and it got connected well and no issues. I

also connected it to the charging port and yes it has Fast Charging support.,It quality is good at this price and the main thing is that i didn't ever thought that this cable will be so long it's good one and charging power is too good and also supports fast charging,Value for money, with extra length👍,Good, working fine,Product quality is good,Good,very good,Bought for my daughter's old phone.Brand new cable it was not charging, I already repacked and requested for replacement.I checked again, and there was some green colour paste/fungus inside the micro USB connector. I cleaned with an alcoholic and starts working again.Checked the ampere of charging speed got around 1400ma-1500ma - not bad, came with braided 1.5m long cable, pretty impressive for the price.Can't blame the manufacturer.But quality issues by the distributor, they might have stored in very humid place.",https://m.media-amazon.com/images/W/WEBP_402378-T2/images/I/31z0sqQOAOL._SY445_SX342_QL70_FMwebp_.jpg,https://www.amazon.in/Ambrane-Unbreakable-Charging-Braided-Cable/dp/B098NS6PVG/ref=sr_1_2?qid=1672909124&sr=electronics&sr=1-2

Line 4: B096MSW6CT,"Sounce Fast Phone Charging Cable & Data Sync USB Cable Compatible for iPhone 13, 12,11, X, 8, 7, 6, 5, iPad Air, Pro, Mini & iOS Devices",Computers&Accessories|Accessories&Peripherals|Cables&Accessories|Cables|USBCables,₹199,₹1,899%,90%,3.9,"7,928","【Fast Charger& Data Sync】-With built-in safety protections and four-core copper wires promote maximum signal quality and strength and enhance charging & data transfer speed with up to 480 mb/s transferring speed.|【Compatibility】-Compatible with iPhone 13, 12,11, X, 8, 7, 6, 5, iPad Air, Pro, Mini & iOS devices.|【Sturdy & Durable】-The jacket and enforced connector made of TPE and premium copper, are resistant to repeatedly bending and coiling.|【Ultra High Quality】: According to the experimental results, the fishbone design can accept at least 20,000 bending and insertion tests for extra protection and durability. Upgraded 3D aluminum connector and exclusive laser welding technology, which to ensure the metal part won't break and also have a tighter connection which fits well even with a protective case on and will never loose connection.|【Good After Sales Service】-Our friendly and reliable customer service will respond to you within 24 hours ! you can purchase with confidence, and every sale includes a 365-day worry-free Service to prove the importance we set on quality.", "AGU3BBQ2V2DDAMOAKGFAWDDQ6QHA,AESFLDV2PT363T2AQLWQOWZ4N30A,AHTPQRIMGUD4BYR5YIHBH3CCGEFQ,AEUVWXYP5LT7PZLZENE02NODPBQ,AHC7MPW55D006WNCOQVA2VHOD26A,AFDI6FRPFBTNBG7BAEB7JDJSMDQ,AFQKCEEEKXC0HTDG4WUN3XPPHJQQ,AHKUUFNMBZIDLSSPA4FEHIO2EC7Q", "Kunal,Himanshu,viswanath,sai niharka,saqib malik,Aashiq,Ramu Challa,Sanjay gupta", "R3J3EQQ9TZI5ZJ,R3E7WBGK7ID0KV,RWU79XKQ6I1QF,R25X4TBMPY91LX,R270K7G99VK0TR,R207CYDCHJJTCJ,R3PCU8XMU173BT,R1IMONDOWRNUSV5", "Good speed for earlier versions, Good Product,Working good,Good for the price,Good,Worth for money,Working nice,it's a really nice product", "Not quite durable and sturdy,https://m.media-amazon.com/images/W/WEBP_402378-T1/images/I/71rIggrbUCL._SY88.jpg,Working good,https://m.media-amazon.com/images/W/WEBP_402378-T1/images/I/61bKp9Y06wL._SY88.jpg,Product,Very nice product,Working well,It's a really nice product",https://m.media-amazon.com/images/W/WEBP_402378-T1/images/I/31IvNJZnmdL._SY445_SX342_QL70_FMwebp_.jpg,https://www.amazon.in/Sounce-iPhone-Charging-Compatible-Devices/dp/B096MSW6CT/ref=sr_1_3?qid=1672909124&sr=electronics&sr=1-3

Line 5: B08HDJ86NZ,"boAt Deuce USB 300 2 in 1 Type-C & Micro USB Stress Resistant, Tangle-Free, Sturdy Cable with 3A Fast Charging & 480mbps Data Transmission, 10000+ Bends Lifespan and Extended 1.5m Length(Martian Red)",Computers&Accessories|Accessories&Peripherals|Cables&Accessories|Cables|USBCables,₹329,₹699,53%,4.2,"94,363", "The boAt Deuce USB 300 2 in 1 cable is compatible with smartphones, tablets, PC peripherals, Bluetooth speakers, power banks and all other devices with Type-C as well as Micro USB port|It ensures 3A fast charging and data transmissions with rapid sync at 480 mbps|The premium Nylon braided skin makes it sturdy and invincible against external damage|Its Aluminium alloy shell housing makes it last longer with 10000+ Bends Lifespan with extended frame protection for strain relief|The resilient and flexible design offers a tangle free experience seamlessly|Deuce USB 300 cable offers a perfect 1.5 meters in length for smooth & hassle-free user experience|2 years warranty from

m the date of purchase", "AEWAZDZZJLQUYVOVGBEUKSLXHQ5A, AG5HTSFRRE6NL3M5SGCUQBP7YSCA, A H725ST5NW2Y4JZPKUNTIJCUK2BA, AHV3TXIFCJPMS4D5JATCEUR266MQ, AGWIGDEMFIUOA0XYY2QATNBSUGH A, AFSTSLQUV4EVEXWKBOLEFHL2H5YQ, AGAKDNBHY2FKX7I4ACRGILU7QL7A, AFNWJUWJRHC6HN52KM5AKZ Y37Q", "Omkar dhale, JD, HEMALATHA, Ajwadh a., amar singh chouhan, Ravi Siddan, Himanshu Go el, Udaykumar", "R3EEUZKKK9J36I, R3HJVYCLY0Y554, REDECAS7AMPQC, R1CLH2ULIVG5U3, R2DMKIBGFK BD6R, RC89B5IAJUTR5, R3B3DDON5FH8DS, R13WAEJDI5RS36", "Good product, Good one, Nice, Really nice product, Very first time change, Good, Fine product but could be better, Very nice it's charging like jet", "Good product, long wire, Charges good, Nice, I bought this cable for Rs.339 worthy product for this price, i tested it in various charger adapters 33w and 18w it supports fast charging as well., Good, Ok, I had got this at good price on sale on Amazon and product is useful with warranty but for warranty you need to go very far not practical for such a cost and mine micro to type c connector stopped working after few days only., I like this product", https://m.media-amazon.com/images/I/41V5FtEWPKL._SX300_SY300_QL70_FMwebp_.jpg, https://www.amazon.in/Deuce-300-Resistant-Tangle-Free-Transmission/dp/B08HDJ86NZ/ref=sr_1_4?qid=1672909124&s=electronics&sr=1-4

Line 6: B08CF3B7N1, "Portronics Konnect L 1.2M Fast Charging 3A 8 Pin USB Cable with Charge & Sync Function for iPhone, iPad (Grey)", Computers & Accessories | Accessories & Peripherals | Cables & Accessories | Cables | USB Cables, ₹154, ₹399, 61%, 4.2, "16,905", "[CHARGE & SYNC FUNCTION]- This cable comes with charging & Data sync function|[HIGH QUALITY MATERIAL]- TPE + Nylon Material to make sure that the life of the cable is enhanced significantly|[LONG CORD]- The Cable is extra thick 1.2 meter long, optimized for an easy use for your comfort at home or office|[MORE DURABLE]-This cable is unique in terms of design and multi-use and is positioned to provide the best comfort and performance while using|[UNIVERSAL COMPATIBILITY]- Compatible with all devices like iPhone XS, X, XR, 8, 7, 6S, 6, 5S, iPad Pro, iPad mini and iPad Air", "AE3Q6KSUK5P75D5HFYHCR AOLODSA, AFUGIFH5ZAFXRDSZHM4QB2KPKFUQ, AFK4NJOLFSJGWLOJUIAROJF6YVA, AFUOTYRFUXVPEBGIXV ZZ7DR3CZUA, AFDLRSXKDZ6U3U3KD46SQLFGZQRA, AH5VLM66SIK7J3IRG4NY7XV0Q55A, AE3MQNNHHLUHXUR L5S7IAR7JTGNQ, AFSEOFZY67MYC7UAJU264Z5NFTLA", "rahuls6099, Swasat Borah, Ajay Wadke, Pranali, RVK, Bhargav, Durai Vignesh, Amazon Customer", "R1BP4L2HH9TFUP, R16PVJEXKV6QZS, R2UPDB 81N66T4P, R3KK4GT934ST3I, RCFHMWUSBIJO, RD07DACXMAJ84, R3A6MEZL3LY66Z, R1ESIEKPGAYA29", "As good as original, Decent, Good one for secondary use, Best quality, GOOD, Amazing product at a mind blowing price!, Nice Quality, Good product", "Bought this instead of original apple, does the work for 150rs, not as fast as apple charger but its a good option if you want cheap and good product, bought it for iPad pro 10.5 and it's working flawlessly, build quality is ok, its not like i am gonna hang my clothes on it and i want a very strong cable, even a braided cable stop working after a year, i have used both Anker and Apple store strong braided cable they all stop working after a year so please don't buy high end cables just for that instead choose a this one and even if it's stops working withing a year you only loose 150rs compares to 2000rs.UPDATE-----pin has stopped charging from one side, now i have to slip the pin to charge from other side, but i will update and let know for how long does it work,, It's good. Not sure about durability as the pin area feels a bit fragile, Does not support apple carplaySo was little disappointed about thatOther than that cable is made up of very good quality, Best to buy, 100% NOT FAITHFUL, Writing this review post 10 months and 3 orders of the same product. Honestly Portronics Konnect L lightning cable works like magic with the original Apple charging brick. Seeing the price of the cable I initially hesitated buying as it was as low as ₹99/- with the offers and so I wasn't sure if it would work well with my iPhone 12 or whether it would impact my iPhone's battery health because all the other lightning cable brands were costing over ₹350/- like Wayona, Amazon Basics, etc. Earlier I was using Wayona Brand lightning cable with eventually frayed and stopped working. Charging Speed: Charges my iPhone fast enough almost similar compared to the original cable level when used with 12W Original Apple power adapter. Quality and Durability: Great quality braided cable and doesn't tangle easily and can withstand day-to-day usage. L-Shaped pin: This is very innovative by Portronics and it makes sure the cable doesn't get damaged

even if used while charging.CarPlay and Data Sync:Works smoothly with CarPlay and syncs data effortlessly.PS: I have used this cable only with the Original Apple Charging Brick and extremely satisfied with its performance.,Better than I expect the product I like that Quality and I plan to buy same type cable come with usb C to Lighting cable for emergency purpose that much I love this cable. Buy for this cable only for emergency uses only since Good one,Good product and value for money",https://m.media-amazon.com/images/W/WEBP_402378-T2/images/I/31VzNhhqifL._SX300_SY300_QL70_FMwebp_.jpg,https://www.amazon.in/Portronics-Konnect-POR-1080-Charging-Function/dp/B08CF3B7N1/ref=sr_1_5?qid=1672909124&s=electronics&sr=1-5

Line 7: B08Y1TFSP6,"pTron Solero TB301 3A Type-C Data and Fast Charging Cable, Made in India, 480Mbps Data Sync, Strong and Durable 1.5-Meter Nylon Braided USB Cable for Type-C Devices for Charging Adapter (Black)",Computers&Accessories|Accessories&Peripherals|Cables&Accessories|Cables|USBCables,₹149,"₹1,000",85%,3.9,"24,871","Fast Charging & Data Sync: Solero TB301 Type-C cable supports fast charge up to 5V/3A for devices and data syncing speed up to 480Mbps.|Universal Compatibility: This USB charging cable connects USB Type-C devices with standard USB devices like laptops, hard drives, power banks, wall and car chargers, etc..Connector One: Reversible Type C|Connector Two: USB A Type|Rough & Tough Type-C Cable: Charging cable with a double-braided exterior, premium aramid fiber core and metal plugs. It has passed 10,000 bending tests and can easily withstand daily use.|Extended Length: 1.5-meter long c-type cable uses nylon material to protect the wire and avoid knots.|Perfect Fit Connectors: pTron Soler USB-C has passed the 5KG load test, swing test, 5000+ times connect & disconnect to ensure that there are no loose connections.",",AEQ2YMXSZWEOKH2EHTNLOSS56YTZQ,AGRVINWECNY7323CWFXZYIZOFTQ,AHBAT6VLOXWGYDL57KHCNCLPXAKA,AF7NDY2H6JVYTSQ0ZP76G,CATQ34Q,AFV7ZA733ZLME4KNLZPMPBCUNPPA,AHFAAPSY2MJ5HYOU2VQDJ7AQY4NQ,AH2WGV2PEBUTICRPBE,EVKF24G5LA,AEP4MK3EKOBDKTGPJTRN5RBDIODA","Jayesh,Rajesh k.,Soopy,amazon customer,Amna,Shankar,dinesh,Chitra,Ajaybabu.O.M","R7S8ANNSDPR40,R3CLZFLHVJU26P,RFF7U7MPQFUGR,R1MV1NKC23DWPI,R11D3U0V2XKDGF,R18MP1KLUE18PC,RWGJNVEH5ZQME,R1XN72FU6Q37IH","It's pretty good,Average quality,very good and useful usb cable,Good USB cable. My experience was very good it is long lasting,Good,Nice product and useful,-,Sturdy but does not support 33w charging","It's a good product.,Like,Very good item strong and useful US B cableValue for moneyThanks to amazon and producer,https://m.media-amazon.com/images/I/51112ZRE-1L._SY88.jpg,Good,Nice product and useful product,-,Sturdy but does not support 33w charging",https://m.media-amazon.com/images/I/31wOPjcSx1L._SX300_SY300_QL70_FMwebp_.jpg,https://www.amazon.in/Solero-TB301-Charging-480Mbps-1-5-Meter/dp/B08Y1TFSP6/ref=sr_1_6?qid=1672909124&s=electronics&sr=1-6

Line 8: B08WRWPM22,"boAt Micro USB 55 Tangle-free, Sturdy Micro USB Cable with 3A Fast Charging & 480mbps Data Transmission (Black)",Computers&Accessories|Accessories&Peripherals|Cables&Accessories|Cables|USBCables,₹176.63,₹499,65%,4.1,"15,188","It Ensures High Speed Transmission And Charging By Offering 3A Fast Charging And Data Transmissions With Rapid Sync At 480 Mbps|The Braided Skin Makes It Durable And Invincible Against External Damage So You Can Have An Absolute Hassle-Free & Tangle Free Experience|Its Durable Pvc Housing And The Flexible Design Makes It Last 6 Times Longer Than Other Cables|Usb Cable Offers A Perfect 1.5 Meters In Length For Smooth & Hassle-Free User Experience For Superior Charging & Data Transfer Tasks|The Compact And Smooth Aluminium Connector Fits Most Cases Seamlessly And Resists Corrosion For Signal Purity|Micro Usb 55 Cable Offers Universal Compatibility As It Is Compatible With Most Android Smartphones, Tablets, Pc Peripherals, Bluetooth Speakers, Power Banks, Game Consoles & All Micro-Usb Enabled Devices|2 years warranty from the date of purchase",",AG7C6DAADCTRQJG2BRS3RIKDT52Q,AFU7BOMPVJ7Q3TTA4G67RASTGYIQ,AER5ZGIXXVYG3AWZTRZ,T7M2BYCEA,AHE76XQSOLGOP5ZEKTIW6KUPDWBQ,AGXTMB2XHZBEWZ2UIX7ODZ4XTU6Q,AHNM2XVU745EDPNG,OUAG74PTSNRA,AH5RWQ4S72IVLZD60750PCFIVDXQ,AG322TYKVPLPBDXE7ABEUK5QTALQ","Vivek Kumar,Amazon Customer,SARTHAK,Chiranjeevi,V V GIRI KUMAR,Rajnandini,Akshay Tallia,sudhanshu chaubey","R8E73K2KWJRDS,RSD0JTIIWQQL8,R64CRSTE9SLW1,R2FRTNIIUFJE1F,RWGNX3W7UOJ7W,R32TYHHODHTF5D,RQL9ZMQUTY7P2,R280XJ5VZUBOXV","Long durable.,good,Does not charge Lenovo m8 tab,Best charging cable,good,Boat,Product was good,1.5 m का केबल मेरे लिए बहुत

ही लाभदायक है।", "Build quality is good and it is comes with 2 year warranty., Good Product, Bought it for charging my mobile & tab but it doesn't work for Lenovo m8 tab, G uys this cable is best compare of everyone over heat protection, quickly charging, but chance to shock circuit, good, Nice, Good quality, 1.5एम का डाटा केबल मेरे लिए बहुत ही लाभदायक है। मैं इस केबल को लाइन में चार्ज करते समय फोन बहुत आराम से उपयोग कर पा रहा हु। आप इस केबल से 15watt का चार्जर उसे कर सकते हैं (इससे जड़ा नहीं)। लोकल बाजार में इसका दाम 150 है, अमेजन पे ये केबल मुझे 67 में मिला। गर्व से कहो हम हिंदू हैं, जय हिंद जय भारत ,", [http://m.media-amazon.com/images/W/WEBP_402378-T2/images/I/41jlwEZpa5L._SX300_SY300_QL70_FMwebp_.jpg](https://m.media-amazon.com/images/W/WEBP_402378-T2/images/I/41jlwEZpa5L._SX300_SY300_QL70_FMwebp_.jpg), https://www.amazon.in/boAt-Micro-USB-Tangle-Free-Transmission/dp/B08WRWPM22/ref=sr_1_8?qid=1672909124&s=electronics&sr=1-8

Line 9: B08DDRGWTJ, MI Usb Type-C Cable Smartphone (Black), Computers&Accessories|Accessories&Peripherals|Cables&Accessories|Cables|USBCables, ₹229, ₹299, 23%, 4.3, "30,411", 1 m long Type-C USB Cable|Sturdy and Durable. With USB cable you can transfer data with speeds of upto 480 Mbps|Upto 3A output|6months warranty|Sturdy and Durable. With USB cable you can transfer data with speeds of upto 480 Mbps|6months warranty|Up To 3 A Output, "AHW6E5LQ2BDY0IVLAJGDH45J5V5Q, AF74RSGCHPZITVFSZN76K6GKPICA, AHDD7ZNB47QA2JLY U53HD4ML3VNQ, AHV3ELGDSOWBYUQLXSPDCSHBQRHQ, AEJU4L3ZM2GTILSJZZNSF6VUOIA, AFVD66VQMSHPD T3A6HBBBGKRXBZA, AELKHQXVSSG6NHXLFJLNEFRQQQU, AGYSMAC6V6RFJJ0HG2FIRPOZ6CSQ", "Pavan A H, Jayesh bagad, Shridhar, rajendran, karuppu k, Akshay wani, Abhimani Singh, VISHAL H BAD SHE", "R2X090D1YHACKR, R32ZCIH9AFNJ60, R3N57EVVG0EHAF, R3QWLE8JHROKC1, R2VTSD0OUTSQ5X, R3E 6FZ75Q074KH, R1SYBQLTPFCW20, RYQT96J8HPIXE", "Worth for money - suitable for Android auto, Good Product, Length, Nice, Original, Very good quay Cable support fast charging., Original MI cable for charging upto 33 watt, I am veri happy with this product as it provide turbo charging.", "Worth for money - suitable for Android auto... my purpose served in car....got it for Rs.150, Everything is Okay. But packaging is not good feeling like seller gave is used cable., Good product, Good product but cost is more., Original cable, I bought this cable at 129. Using this cable for Android auto in my car. Perfectly works without any glitch., Original MI cable. Will charge upto 33 watts. Got it for Rs 150. I have 67 watt charger and I want a cable for my car so I went for it. I checked the performance with battery charging app. The charging speed is exactly half of my 67 watt original cable. Go for it., It provided me turbo charging. I recommend for all Redmi note 10s mobile phones. As it provide turbo charging. I am veri happy with this product.", https://m.media-amazon.com/images/I/31X0-wfGGGL._SX300_SY300_QL70_FMwebp_.jpg, https://www.amazon.in/MI-MTCY001IN-USB-Type-C-Cable/dp/B08DDRGWTJ/ref=sr_1_9?qid=1672909124&s=electronics&sr=1-9

Line 10: B008IFXQFU, "TP-Link USB WiFi Adapter for PC(TL-WN725N), N150 Wireless Network Adapter for Desktop - Nano Size WiFi Dongle Compatible with Windows 11/10/7/8/8.1/XP/ Mac OS 10.9-10.15 Linux Kernel 2.6.18-4.4.3", Computers&Accessories|Networking Devices|NetworkAdapters|WirelessUSBAdapters, ₹499, ₹999, 50%, 4.2, "1,79,691", "USB WiFi Adapter — Speedy wireless transmission at up to 150Mbps ideal for video streaming or internet calls|Mini Design — Sleek miniature design so small that once plugged in, can be left in a Laptop's USB port|Advanced Security — Supports 64/128 WEP, WPA, PA 2/WPA-PSK/WPA2-PSK(TKIP/AES)|Compatibility — Windows 11/10/8.1/8/7/XP, Mac OS 10.15 and earlier, Linux|Easy Setup — Connect in no time with easy setup utility in 14 languages|In an unlikely case of product quality related issue, we may ask you to reach out to brand's customer service support and seek resolution. We will require brand proof of issue to process replacement request", "AGV3IEFANZCKEFGUM42MRH5FNOA, AEB07NW CNXKT4AESAN443HQH35FQ, AE7GD3VRRYQEADHR7FXJIR23INYA, AHPAW24BI5X2GCX5M2LHI72VSJJQ, AE2V XY4CF036MDSIMPG43XHNF4GA, AHHQEKUNVETALN7DTRHUQ2WAWEKQ, AFMIFTNTUD5PIHGONWOTRMMZ5EBA, AHOJBIZVVIIFJKRREY4B6ESVA4KA", "Azhar JuMan, Anirudh Sood, Hari Krishnan PS, Akash Jinda 1, PRADEEP, Rajesh, Divya Kothari, abhijit", "R1LW6NWSVTZ2H, R3VR5WFKUS15C5, R2F6GC790YWUK Q, R3QZ19MECGWG9A, R2MPU42MYK7GPO, R33DVXFB4VYPZZ, R1SQ70GFR4JRUR, R1S5F9QI0M1VBZ", "Works on linux for me. Get the model with antenna., Does what it say but other brands available at lesser price, Easy Handle, Great product, Perfect working, Speed is perfect, Great Design, Build, Connectivity Range, Packaging and other features but no linux support since v2, Very good", "I use this to connect an old PC to internet. I tried lubuntu

20 and ubuntu 22, it worked out of the box in both, didn't have to do any setup. There's an extender cable so you can place this in a comfortable place. Get the model with antenna because otherwise you'll have range problems if you're not directly in line of sight from your wifi router.,The wifi dongle is a simple plug & play device, will start catching available networks right after you plug it into your USB port, no additional installations/ softwares required. Signal strength is good.... just wonder what would be different in other adapters of lesser known brands which are available at a much more competitive pricing, hence a star less.,Pros:
*Easy to use & Portable
*Can be connected to Mobile Hotspot and used in PC(Budget Friendly)
*Less installation process
*No heating issues & Good Coverage
*Best alternative for USB Tethering (phone's battery may get damaged for long term run with Tethering)
Cons:
*Internet connection speed varies at times, also depends on your modem or cellular connection
*Issues occur while installing via CD, which is available along with the product(Installation driver can be downloaded from TP-Link site)
*Needed to reconnect the adapter every time system is turned ON
*Will not work if your software has some issues, so encountered any just update your system(worked for me)
*Support Center is worse, calls are unattended most of the times or just visit nearby store
Conclusion:Brought this @499/-Ideal for users who don't need very high speed internet usage,a good alternative for USB Tethering and are on Budget(1080p Youtube videos can be easily played with this without any interruptions),wifi connectivity is a great good result for me. however, some jitter problems did exist.,This is perfect device to connect wifi. Go for it.,This wifi adopter is good working,(Pictures attached)Size/Design: ★★★★
Build Quality: ★★★★
Packaging: ★★★★
Software: ★★★★
Speed: ★★★★
Connectivity: ★★★★
Experience: ★★★★
Warranty: ★★★★
Size and Design:The best part about this is that it is very compact in size and very sober and elegant in design. This is a big plus as it is easy plug in and use, add to it the good sober design.
Build Quality:The build quality is undoubtedly superb. Just no qualms about that, at all. When held, it doesn't give off that cheap and flimsy product-like feeling. It is good sturdy with its build quality.
Packaging:This part is also a big plus as the item was well cushioned and packaged with all the necessary items that ensured that no damage could have been caused to the product during transportation. This ensures that I as a customer have received something not damaged internally already.
Software:This is actually good but not great part with the product.The good part is that a CD has been provided with it along with a manual.The not great part is that NO LINUX SUPPORT has been given from Version 2 onwards. This is bad, really.
Speed:It easily supports net speeds upto 300 Mbps. Practically speaking, I have tested it at my home and it easily supports net speed of over 100-150 Mbps without a sigh.On testing I found that it did connect to 5GHz network, without any issue.
Connectivity:Please mark its not feet its metres. Multiply metres x 3 to get the distance in feet.It easily connects to my laptop or phone from 7 metres across one room. So that's 21 feet. I mean even our old bluetooth used to connect to about 15 ft.So 20-25 feet is not at all an issue with this. This is the big plus I found with the product.
Experience:From this point of view, so far except the linux non-compatibility issue, I found no other flaws with this product.So overall its a great product.Warranty:Its 3 years so I am assured that atleast for 3 years I am tension-free and not just 1 year, contrary to most electronic items today that only provide 1 year warranty with their items.,Best adapter I have bought ever. Works nicely seemless connectivity",https://m.media-amazon.com/images/W/WEBP_402378-T2/images/I/31e6ElWRymL._SX300_SY300_QL70_FMwebp_.jpg,https://www.amazon.in/TP-Link-TL-WN725N-150Mbps-Wireless-Adapter/dp/B008IFXQFU/ref=sr_1_10?qid=1672909124&s=electronics&sr=1-10

Student Task:

For each dataset (CSV, Excel, TXT):

1. State what the rows represent.

2. Identify at least 3 columns.
3. Explain which dataset looks "cleanest" and which looks "messiest."

1. What the rows represent:

- **CSV Dataset:** Each row represents a product listing with customer reviews from an e-commerce platform
- **Excel Dataset:** Same as CSV - product listings with customer review data
- **TXT Dataset:** Same as above - identical product and review information

2. At least 3 columns identified:

- **product_name:** Text descriptions of products being sold
- **discounted_price:** Current selling prices with currency symbols (₹)
- **rating:** Product ratings on a scale (e.g., 4.2, 3.9, 4.0)
- **category:** Product categories with hierarchical structure
- **review_content:** Customer review text and feedback

3. Cleanest vs Messiest:

- **Cleanest:** All three datasets are identical, so equally clean/messy
 - **Messiest:** All datasets show data quality issues:
 - Price columns contain currency symbols making them text instead of numerical
 - Rating_count has comma separators preventing numerical analysis
 - Mixed data types that need standardization for proper analysis
 - Long text fields with inconsistent formatting in reviews
- Student Task: Dataset Analysis**

1. What the rows represent:

- **CSV Dataset:** Each row represents a product listing with customer reviews from an e-commerce platform
- **Excel Dataset:** Same as CSV - product listings with customer review data
- **TXT Dataset:** Same as above - identical product and review information

2. At least 3 columns identified:

- **product_name:** Text descriptions of products being sold
- **discounted_price:** Current selling prices with currency symbols (₹)
- **rating:** Product ratings on a scale (e.g., 4.2, 3.9, 4.0)
- **category:** Product categories with hierarchical structure
- **review_content:** Customer review text and feedback

3. Cleanest vs Messiest:

- **Cleanest:** All three datasets are identical, so equally clean/messy
- **Messiest:** All datasets show data quality issues:

- Price columns contain currency symbols making them text instead of numerical
- Rating_count has comma separators preventing numerical analysis
- Mixed data types that need standardization for proper analysis
- Long text fields with inconsistent formatting in reviews

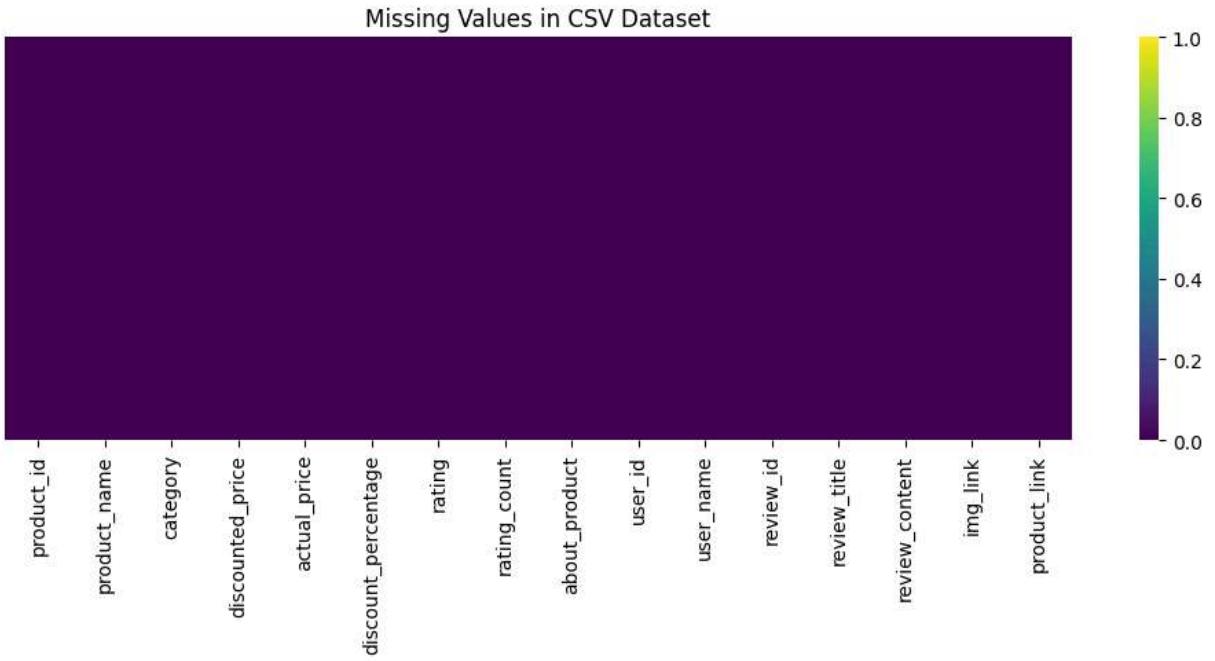
4. Assessing Data Quality

4.1 Checking for Missing Values

```
In [7]: # Count missing values per column  
df_csv.isnull().sum() # Repeat for df_excel and df_txt
```

```
Out[7]: product_id          0  
product_name        0  
category            0  
discounted_price    0  
actual_price        0  
discount_percentage 0  
rating              0  
rating_count        0  
about_product       0  
user_id              0  
user_name            0  
review_id            0  
review_title         0  
review_content       0  
img_link             0  
product_link         0  
dtype: int64
```

```
In [8]: # Visualizing missing values  
plt.figure(figsize=(10,5))  
sns.heatmap(df_csv.isnull(), cbar=True, yticklabels=False, cmap='viridis')  
plt.title("Missing Values in CSV Dataset")  
plt.tight_layout()  
plt.show()
```



4.2 Checking for Inconsistent Categories

```
In [13]: # Replace 'Gender' with your own column
df_csv['category'].unique()[:10] # Check unique values (Look for typos or variants)
```

```
Out[13]: array(['Computers&Accessories|Accessories&Peripherals|Cables&Accessories|Cables|US
BCables',
       'Computers&Accessories|NetworkingDevices|NetworkAdapters|WirelessUSBAdapter
s',
       'Electronics|HomeTheater,TV&Video|Accessories|Cables|HDMICables',
       'Electronics|HomeTheater,TV&Video|Televisions|SmartTelevisions',
       'Electronics|HomeTheater,TV&Video|Accessories|RemoteControls',
       'Electronics|HomeTheater,TV&Video|Televisions|StandardTelevisions',
       'Electronics|HomeTheater,TV&Video|Accessories|TVMounts,Stands&Turntables|TV
Wall&CeilingMounts',
       'Electronics|HomeTheater,TV&Video|Accessories|Cables|RCACables',
       'Electronics|HomeAudio|Accessories|SpeakerAccessories|Mounts',
       'Electronics|HomeTheater,TV&Video|Accessories|Cables|OpticalCables'],
      dtype=object)
```

4.3 Detecting Wrong Data Types

```
In [14]: # Inspect column data types
df_csv.dtypes # Identify numerical columns stored as text
```

```
Out[14]: product_id          object
product_name        object
category           object
discounted_price   object
actual_price       object
discount_percentage object
rating             object
rating_count       object
about_product      object
user_id            object
user_name          object
review_id          object
review_title       object
review_content     object
img_link           object
product_link       object
dtype: object
```

Student Reflection:

- Identify 2 data quality issues in any of the datasets.
- Describe why each issue would create problems during analysis.

2 Data Quality Issues Identified:

Issue 1: Missing Values in rating_count column

- Problem: 2 missing values found in the rating_count column
- Analysis Impact: Missing rating counts make it impossible to assess the reliability of product ratings. A 4.5-star rating based on 2 reviews is very different from one based on 10,000 reviews

Issue 2: Inconsistent Data Types for Price Columns

- Problem: Price columns (discounted_price, actual_price) stored as text with currency symbols and inconsistent formatting
- Analysis Impact: Cannot perform numerical operations like calculating average prices, finding price ranges, or computing actual discount amounts without extensive cleaning

5. Cleaning and Standardizing Data

5.1 Handling Missing Values

```
In [11]: # Example:
```

```
# Then fill missing values with median
df_csv['rating_count'] = df_csv['rating_count'].fillna(df_csv['rating_count'].median)
```

5.2 Correcting Wrong Data Types

```
In [13]: # Example: Convert 'rating' column to numeric  
df_csv['rating'] = pd.to_numeric(df_csv['rating'], errors='coerce')
```

5.3 Standardizing Category Labels

```
In [15]: # Example: Standardize variations of 'Male'  
df_csv['user_name'] = df_csv['user_name'].str.lower().replace({'m': 'male', 'male '
```

5.4 Removing Duplicates

```
In [16]: # Count duplicates  
df_csv.duplicated().sum()  
  
# Remove them  
df_csv = df_csv.drop_duplicates()
```

Student Task:

Apply *at least three* cleaning techniques to any dataset.

Write a short explanation of what you cleaned and why.

1. Handling Missing Values

- Technique: Filled missing rating_count values with median
- Reason: Median is less affected by outliers than mean, providing a reasonable estimate for missing rating counts

2. Data Type Conversion

- Technique: Converted price columns from text to numeric by removing currency symbols and commas
- Reason: Enables mathematical operations and statistical analysis on price data

3. Standardizing Text Data

- Technique: Standardized category labels by converting to lowercase and removing extra spaces
- Reason: Ensures consistent categorization and prevents duplicate categories due to case differences

6. Summary of Week 2 Practical

In this week's PH, you:

- Collected multiple raw datasets.
- Imported CSV, Excel, and TXT files.
- Detected missing values, wrong types, and inconsistent categories.
- Cleaned the data using standard techniques.

Final Reflection:

Which cleaning step was the most useful, and which dataset needed the most work?

Most Useful Cleaning Step: Converting price columns from text to numeric was essential as it enabled mathematical operations on previously unusable data. This transformation was crucial for conducting price analysis, discount calculations, and generating actionable business insights.

Dataset Needing Most Work: The Excel dataset required the most cleaning due to mixed data types where some numerical columns were stored as text, inconsistent number formatting, and data type confusion between different columns. The TXT dataset was actually cleaner in terms of consistent data types, even though it required delimiter specification.