**DATA SCIENCE**

1. **TOOLS FOR DATA SCIENCE**

**1. EXCEL**

**SOME BASIC POINTS**

1. ABC=COLUMN, 123=ROW, CELL, SHEET
2. SYNTAX MEANS (=)
3. RANGE
4. FILLING
5. PARENTHESES MEANS BRACKET

**SOME EXCEL COMMON FUNCTIONS :-**

1. **IF - IFS -** =IF(B1="GRASS","YES","NO") =IFS(C2>90,"FAST",C2>50,"NORMAL",C2<=50,"SLOW")
2. **(AND - OR) (IF - AND - OR) -** =AND(B2="FIRE",C2>70)=IF(AND(B2="FIRE",C2>70),"GOOD","BAD") =OR(B2="FIRE",C2>=70)
3. **SUM-SUMIF-SUMIFS** - =SUM(K26:K34) SUMIF - =SUMIF(RANGE, CRITERIA, [SUM\_RANGE]) SUMIFS - =SUMIFS(SUM\_RANGE, CRITERIA\_RANGE1, CRITERIA1, [CRITERIA\_RANGE2, CRITERIA2] ...)
4. **AVERAGE - AVERAGEIF -** AVERAGEIFS =AVERAGE(RANGE) =AVERAGEIF(RANGE, CRITERIA, [AVERAGE\_RANGE]) =AVERAGEIFS(AVERAGE\_RANGE, CRITERIA\_RANGE1, CRITERIA1, ...)
5. **COUNT- COUNTA - COUNTBLANK - COUNTIF** - COUNTIFS =COUNT(K26:K34) =COUNTA(I26:I34) =COUNTBLANK(K45:K55) =COUNTIF(RANGE,CRITERIA) =COUNTIFS(CRITERIA\_RANGE1, CRITERIA1, [CRITERIA\_RANGE2, CRITERIA2], ...)
6. **LOWER - UPPER** =LOWER(A41:A47) =UPPER(A41:A47)
7. MAX-MIN - FINDS THE HIGHEST NUMBER IN A RANGE =MAX(K26:K34) =MIN(K26:K35)
8. **MEDIAN-MODE** - =MEDIAN(K26:K34) =MODE.SNGL(B42:F47)
9. **NPV -**NET PRESENT VALUE=NPV(RATE, VALUE1, VALUE2, ...)
10. **LEFT-RIGHT-**.=LEFT(A95:A97,4) EXTRACT RIGHT DIGITS. =RIGHT(A95:A97,3)
11. **STDEV.P -STDEV.S -** =STDEV.P(C95:C114) =STDEV.S(E119:E138)
12. **CONCAT- CONCATENATE-**  =CONCAT(A95," ",A96)
13. **RAND -** =RAND()\*100 , =RAND()
14. **TRIM -** =TRIM(A150:C169)
15. **VLOOKUP -** =VLOOKUP(LOOKUP CELL,TABLE, COLUMN NO, 1/0). IF MOST LEFT TEXT IS NO 1. IF MOST LEFT TEXT IS TEXT 0.
16. **HLOOKUP -** HLOOKUP(ORDER ID,TABLE, COLUMN NO,FALSE) TRUE=APPROXIMATE MATCH, FALSE = PERFECT MATCH
17. **LEN -** =LEN(A2)
18. **DAYS/NETWORKDAYS** - =DAYS(DATE1,DATE2)=NETWORKINGDAYS(SELECT CELL, SELECT CELL,[NUMBEROFHOLIDAYS])
19. **FIND/SEARCH -** =VALUE(LEFT(G2,FIND(“ “,G2))
20. **IFERROR -** =IFERROR(F17/G17,0)
21. **RANK -** =RANK(BASE CEL,CELRANGE)
22. **SUMPRODUCT -** =SUMPRODUCT(PRODUCTPRICERANGE,PRODUCTQUANTITYRANGE) =SUMPRODUCT(PROCUCTRANGE=SPECIFICPRODUCT)\*PRODUCTPRICERANGE\*PRODUCTQUANTITYRANGE
23. **SUBSTITUTE -** =SUBSTITUTE(D4,” “,””)
24. **VALUE -** =VALUE(LEFT(G2,FIND(“ “,G2))
25. **UNIQUE -** =UNIQUE(RANGE)
26. **PERCENTILE -** =PERCENTILE(RANGE,0.09)
27. **QUARTILE -** =QUARTILE(RANGE,1/2/3/4/0)
28. **COVAR-** =
29. **GEOMEAN**

**EXCEL FEATURES FOR DATA ANALYSIS**

1. **HOME CELLS -** UNDO-REDO, CUT-COPY-PASTE, FORMAT PRINTER,FONT, ALLIGNMENT, NUMBERS, **CONDITIONAL FORMATING,** FORMAT AS TABLE, CELL STYLE, INSERT-DELETE-FORMAT CELS, FILL,CLEAR, FIND AND SELECT
2. **INSERT - PIVOT TABLE,** TABLE, INLUSTRATION, ADDINDS, CHARTS, 3D MAP, SPARKLINES,FILTERS, LINK, COMMENT, TEXT, SYMBOL,
3. **PAGE LAYOUT -** THEAMS, PAGE SETUP, SCALE TO FIT, GRIDLINES, HEADINGS, ARRANGE
4. **REVIEW -** SPELLING,THESAURUS,WORKBOOK STATISTICS, CHECK ASSECIBILITY, SMART LOOKUP, TRANSLATE, COMMENTS, NOTE, PROTECT, INK
5. **VIEWS-** WORKBOOK VIEWS,SHOW, ZOOM, WINDOW, MACRO
6. **DATA -** GET AND TRANSFORM DATA,TEXT TO COLUMN, FLASH FILL,REMOVE DUPLICATES,DATA VALIDATION, CONSOLODITATE,RELATIONSHIP, MANAGE DATA MODEL,WHAT IF ANALYSIS,FORECAST SHEET, QUARIES AND CONNECTIONS,GROUP,UPGROUP,SUBTOTAL

SAMPLE EXCEL WITH ALL FUNTIONS CHECK TO BE PREPARED

My Sql, python, statistics, math, git, data science, Machine Learning,