Quintile\_REPORT.xlsx

Quintile

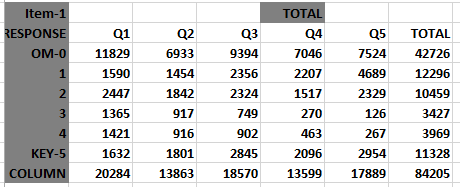
Quintile

Quintile\_REPORT.xlsx

1. Generate responses from TCS data i.e. (Option id’s selected by candidate are of 10 or more digits and it needs to be converted to single digit response.
2. Calculate scores of each candidate. **{(if response selected = final key provided) then add 1 and calculate score of candidate)**}
3. **Total field in each item = Number of candidates who selected that particular response.**

(where responses are :- 0 –for candidates who didn’t select any response.

1. Candidates who selected option 1.
2. Candidates who selected option 2.
3. Candidates who selected option 3.
4. Candidates who selected option 4.
5. Candidates who selected option 5)



1. Calculate **no\_required** i.e. **No of candidates appeared for test/5**
2. Calculate **max\_no** i.e. **No of candidates appeared for test/4.5**
3. Calculate **highest score** among all the candidates.
4. Calculate **cummulative frequency** in decreasing order of highest score.
5. Then find out cutoff for **Q1,Q2,Q3,Q4** by below method :-
6. Q ranges from **1 to 4**
7. If **cummulative\_Frequency < no\_required** then

**Score=Score-1 ;** where k is initialized to maximum number of test

1. Else if **cum\_frequency > max\_no:**

**cut\_off[Q]=Score+1**

1. Else **cut\_off[Q]=Score**

**Score=Score-1**

**Attempted Candidates** :- 84205

**Highest score of candidates**:- 30

**Number Required**= 84205/5 = 16841.0

**Maximum Number** = 18712.2222222

**Sum Frequency Cummulative Frequency**

30 2 2

29 13 15

28 25 40

27 48 88

26 124 212

25 226 438

24 342 780

23 476 1256

22 665 1921

21 916 2837 **Remember,**

20 1170 4007 **Number Required**= 84205/5 = 16841.0

**Maximum Number** = 18712.2222222

19 1342 5349 Here, till score 15 cum\_freq is < no\_required

18 1672 7021 So, by first condition **score=score-1(score=15-1=14)**

17 2035 9056 **2nd condn**:- cum\_freq(14) not > 18712.2222222,

16 2496 11552 Hence score will remain = **14** and **cut\_off(4)=14**

15 2945 14497

14 3392 17889

13 4041 21930

**cut\_off[4] is 14 Next loop will start with cum\_freq of score-1**

13 4041 4041 Start new cumulative frequency for 2nd cut\_off

12 4479 8520 For **Score[11], cum\_freq <** **Number Required**

11 5079 13599 Thus score=score-1 i.e. **score=11-1=10**

10 5733 19332 score[10]>max\_no ,thus score=score+1 **i.e. score=11**

**cut\_off[3] is 11** **Hence, cut\_off[3] is 11**

10 5733 5733 Next Loop starts with cum\_freq of score-1 .Here 10

9 6193 11926

8 6644 18570

7 6975 25545

**cut\_off[2] is 8**

7 6975 6975

6 6888 13863

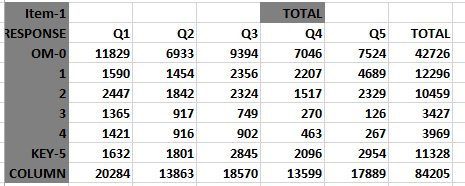
5 6310 20173

4 5388 25561

**cut\_off[1] is 6**

**For Q=1,5**

1. Now for all candidates :-
   1. If **score >= 1** and **score < 6,** those number of candidates will go in Q1
   2. If **score >=6** and **score < 8 ,**those number of candidates will go in Q2
   3. If **score >=8** and **score < 11,** those number of candidates will go in Q3
   4. If **score >=11** and **score < 14,** those number of candidates will go in Q4
   5. If **score >=14,** those number of candidates will go in Q5

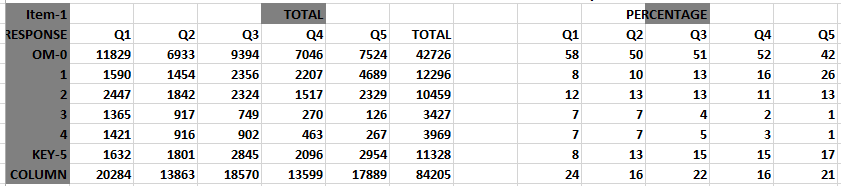


**PERCENTAGE SECTION**

Q1 = respective value in 1st section / (Total of that particular column)

Similarly Q2,Q3,Q4,Q5

(Values printed in Quintile Report are Rounded off)



**STATISTICS**

**MEAN**

1. Calculate Total score of all candidates who selected that particular response for Item.

**Example:-**

candidate 1 respond 0 for item 1 and scored SUM[0] : 7

candidate 2 respond 0 for item 1 and scored SUM[1] : 4

candidate 3 respond 0 for item 1 and scored SUM[2] : 9

candidate 5 respond 0 for item 1 and scored SUM[4] : 11

candidate 6 respond 0 for item 1 and scored SUM[5] : 8

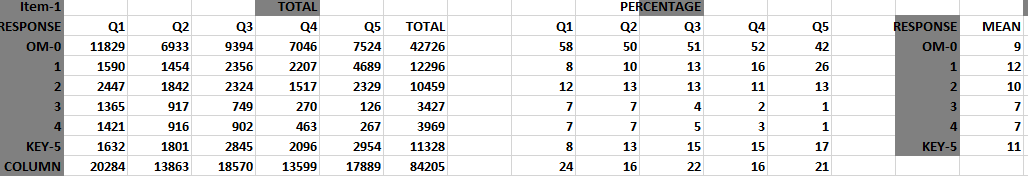
candidate 7 respond 0 for item 1 and scored SUM[6] : 13

candidate 8 respond 0 for item 1 and scored SUM[7] : 4

candidate 9 respond 0 for item 1 and scored SUM[8] : 2

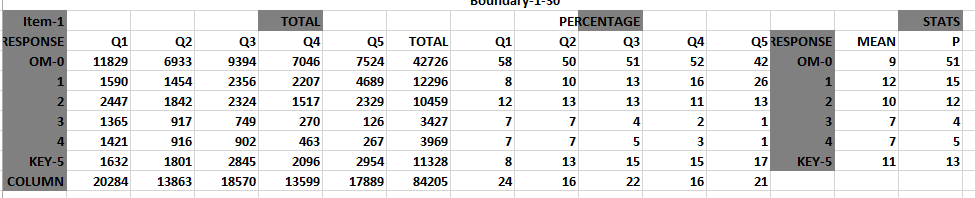
and so on…..

**Tot\_MEAN[0][0] =Sum of above scores/42276 : - 9**



**P value**

1. P = Number of candidate opting the response for ITEM/Candidates who attempted that Item



**R Value**

**R= ((Mp-Mq)/Sd ) \* (SQRT(P\*Q))**

1. Calculate Mean of all candidates for each test.
2. Calculate TOT\_MEAN i.e. Sum of candidates/No of candidates who opted that response.
3. Mexcept= (Mean of all candidates for each test)\* candidates appeared for test -

(Mean calculated above) \*(candidate who selected that particular response)/ (candidate appeared for test - candidate who selected that particular response)

1. P= candidate who selected the particular response/ candidate appeared for test
2. R= (Mean calculated above)-mexcept / Standard Deviation \* sqrt(P\*(1-P))

Hence, Formula used :-

**mexcept**=((((mean(mean\_cand,mean\_add,attempt\_cand,candidate,cnt)))\*float(attempt\_cnd(attempt\_cand,candidate,cnt)))-(TOT\_MEAN[item][opt]\*float(TOT\_ROW[item][opt])))/float((attempt\_cnd(attempt\_cand,candidate,cnt)-TOT\_ROW[item][opt]))

**P**=float(TOT\_ROW[item][opt])/float(attempt\_cnd(attempt\_cand,candidate,cnt))

**R[item][opt]=((**TOT\_MEAN[item][opt])-(mexcept))/(std\_dev\_func())\*math.sqrt((P)\*(1-P))

**SNIPPET**

This worksheet contains summary of all tests. It contains below details :-

1. **Number of candidates who attempted the test.**

(Candidates who did not select any response for whole test are excluded from the total count of candidate)

1. **Highest score of candidates for the test.**

(Candidates scores are compared with each other’s score and highest score is calculated)

1. **Mean of candidates for each test.**

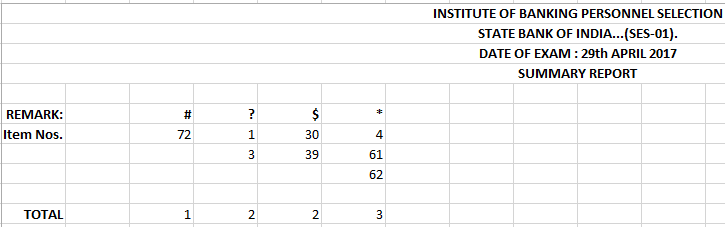
(Total score scored by all candidates for the test / Total number of candidates who attempted the test)

1. **Standard Deviation of candidates for the test.**

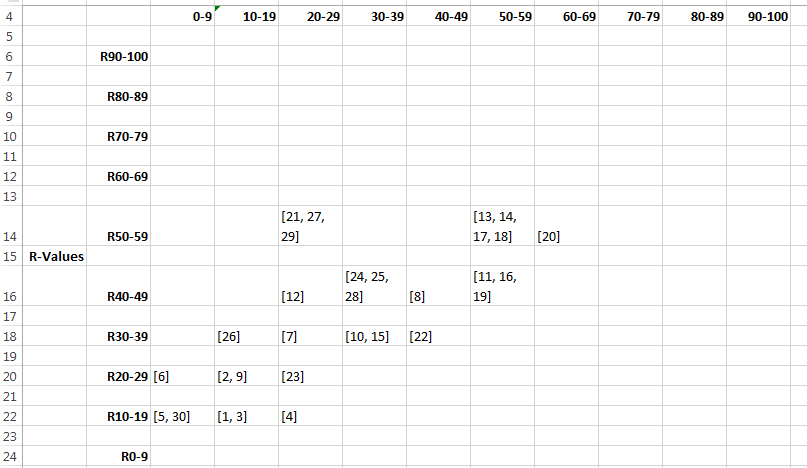
sum=sum+(sum of candidate score for all items)-((sum of all candidate scores for all items)/attemted candidate)

**ITEM\_CHARACTERISTICS.xlsx**

1. **Symbolic Remark**



|  |
| --- |
| **# : Key\_r is negative** **:-** If R value for item for key response is < 0 then that item is marked as “#” . |
|  |
| **? : Key\_r value is less than some other response`s r value :-** IfR value for key response is less than R value of other responses then that item is marked as “?” |
|  |
| **$ : According to highest quintile population the key may be different than population key :- (If Q5(percentage section) for key <= Q5(percentage section) for all reponses then those items will go under $)**  If number of candidates opting for response other than key response is greater than the number of candidates opting for key response then those item are marked under “$” |
|  |
| **\* : Response other than key response has r value greater than or equal to ( Key\_r - 10 ) or 25.**  If R value for other responses of item is greater than R value for key response-10 then that item is marked as “\*”  ! : **Response other than key response has positive R and difference with R value of a key is 4 or more** |
|  |
|  |
| In quintile analysis the r value is calculated using the formulae of Point Biserial Correlation :- A sample size of around 10,000 candidates is usually considered. |
|  |
| The formulae used is as follows : |
| r = ( Mp - Mq ) / sd \* SQRT ( p \* q ) |
| where, |
| Mp = Mean of test scores of candidates who have marked the particular response. |
| Mq = Mean of test scores of candidates who have marked any other response. |
| sd = Standard Deviation of the test scores. |
| p = Percentage population considered for Mp. |
| q = Percentage population considered for Mq.   1. **Shuf.sum**   All items with their symbols is consolidated together in this worksheet.    **P\_R\_ITEM\_PLOT.xlsx** |
| All the items are plotted by their P and R value and worksheets are created for each test. |
|  |



**REPORT\_TO\_ORG.csv**

1. P(other options)  >= P(key)
2. Mean(other options) > = Mean(Key)
3. R(other option) >= 0 and R(other option) >= R(key)
4. R(other option)>= 0 and R(key) – R(other option) > 4