**Medical inventory optimization**

**Exploratory Data Analysis (SQL) by Pratiksha Saheb**

**Software: MySQL Workbench**

**Business decisions based on the clean**

**‘medicines\_db’ data**

**1. Calculating the first moment (measures of central tendency such as mean, median, mode) for the dataset.**

**Mean :**

select

round(avg(Quantity),2) as mean\_Quantity,

round(avg(ReturnQuantity),2) as mean\_ReturnQuantity,

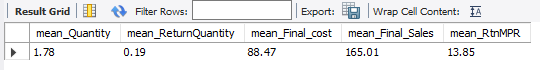
round(avg(Final\_cost),2) as mean\_Final\_cost,

round(avg(Final\_Sales),2) as mean\_Final\_Sales,

round(avg(RtnMPR),2) as mean\_RtnMPR

from

new\_table;



**Median:**

with ranked as

(

select Final\_cost,

Final\_Sales,

Quantity,

ReturnQuantity,

RtnMPR,

row\_number() over (order by Final\_cost) as r,

count(\*) over () as c

from new\_table),

median as

(

select Final\_cost,

Final\_Sales,

Quantity,

ReturnQuantity,

RtnMPR

from ranked

where r in (floor((c+1)/2),ceiling((c+1)/2))

)

select round(avg(Final\_cost), 2),

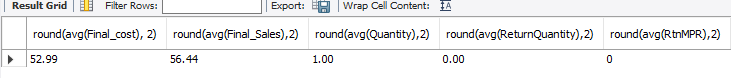
round(avg(Final\_Sales),2),

round(avg(Quantity),2),

round(avg(ReturnQuantity),2),

round(avg(RtnMPR),2)

from median;



**Mode:**

select

mode\_Quantity.mode\_value as mode\_Quantity,

mode\_Quantity.mode\_count as mode\_Quantity\_count,

mode\_ReturnQuantity.mode\_value as mode\_ReturnQuantity,

mode\_ReturnQuantity.mode\_count as mode\_ReturnQuantity\_count,

mode\_Final\_Sales.mode\_value as mode\_Final\_Sales,

mode\_Final\_Sales.mode\_count as mode\_Final\_Sales\_count,

mode\_Final\_cost.mode\_value as mode\_Final\_cost,

mode\_Final\_cost.mode\_count as mode\_Final\_cost\_count,

mode\_RtnMPR.mode\_value as mode\_RtnMPR,

mode\_RtnMPR.mode\_count as mode\_RtnMPR\_count

from(

SELECT Quantity AS mode\_value, COUNT(\*) AS mode\_count

FROM new\_table

GROUP BY Quantity

ORDER BY COUNT(\*) DESC

LIMIT 1 ) as mode\_Quantity,

( SELECT ReturnQuantity AS mode\_value, COUNT(\*) AS mode\_count

FROM new\_table

GROUP BY Quantity

ORDER BY COUNT(\*) DESC

LIMIT 1

) as mode\_ReturnQuantity,

(SELECT final\_cost AS mode\_value, COUNT(\*) AS mode\_count

FROM new\_table

GROUP BY Quantity

ORDER BY COUNT(\*) DESC

LIMIT 1

) as mode\_Final\_cost,

(SELECT Final\_Sales AS mode\_value, COUNT(\*) AS mode\_count

FROM new\_table

GROUP BY Quantity

ORDER BY COUNT(\*) DESC

LIMIT 1

) as mode\_Final\_Sales,

( select RtnMPR AS mode\_value, COUNT(\*) AS mode\_count

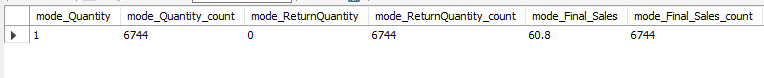
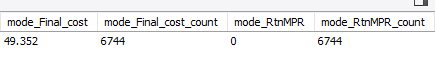
FROM new\_table

GROUP BY Quantity

ORDER BY COUNT(\*) DESC

LIMIT 1

) as mode\_RtnMPR;

**2. Calculating the second moment (measures of dispersion such as variance, standard deviation, range) for the dataset.**

**Variance:**

SELECT

ROUND(VARIANCE(Quantity), 2) AS variance\_quantity,

ROUND(VARIANCE(ReturnQuantity), 2) AS variance\_return\_quantity,

ROUND(VARIANCE(Final\_Cost), 2) AS variance\_final\_cost,

ROUND(VARIANCE(Final\_Sales), 2) AS variance\_final\_sales,

ROUND(VARIANCE(RtnMPR), 2) AS variance\_rtnmrp

FROM new\_table;



SELECT

ROUND(STDDEV(Quantity), 2) AS stddev\_quantity,

ROUND(STDDEV(ReturnQuantity), 2) AS stddev\_return\_quantity,

ROUND(STDDEV(Final\_Cost), 2) AS stddev\_final\_cost,

ROUND(STDDEV(Final\_Sales), 2) AS stddev\_final\_sales,

ROUND(STDDEV(RtnMPR), 2) AS stddev\_rtnmrp

FROM new\_table;

****

**Range:**

SELECT

MAX(Quantity) - MIN(Quantity) AS range\_quantity,

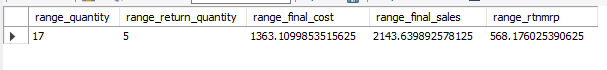
MAX(ReturnQuantity) - MIN(ReturnQuantity) AS range\_return\_quantity,

MAX(Final\_Cost) - MIN(Final\_Cost) AS range\_final\_cost,

MAX(Final\_Sales) - MIN(Final\_Sales) AS range\_final\_sales,

MAX(RtnMPR) - MIN(RtnMPR) AS range\_rtnmrp

FROM new\_table;



**3. Calculating the third moment (skewness) for the dataset.**

**Skewness:**

select Quantity\_skewness1.Quantity\_skewness,

ReturnQuantity\_skewness1.ReturnQuantity\_skewness,

Final\_cost\_skewness1.Final\_cost\_skewness,

Final\_Sales\_skewness1.Final\_Sales\_skewness,

RtnMPR\_skewness1.RtnMPR\_skewness

from

(SELECT

ROUND((SUM(POW(Quantity - (SELECT AVG(Quantity) FROM new\_table), 3)) / (COUNT(\*) \*

POW(STDDEV(Quantity), 3))), 2) AS Quantity\_skewness

FROM new\_table) Quantity\_skewness1,

(select

round((sum(pow(ReturnQuantity - (select avg(ReturnQuantity) from new\_table),3))/(count(\*) \*

pow(stddev(ReturnQuantity),3))),2) as ReturnQuantity\_skewness from new\_table)

ReturnQuantity\_skewness1,

(select

round((sum(pow(Final\_cost - (select avg(Final\_cost) from new\_table),3))/(count(\*) \*

pow(stddev(Final\_cost),3))),2) as Final\_cost\_skewness from new\_table)

Final\_cost\_skewness1,

(

select

round((sum(pow(Final\_Sales - (select avg(Final\_Sales) from new\_table),3))/(count(\*) \*

pow(stddev(Final\_Sales),3))),2) as Final\_Sales\_skewness from new\_table)

Final\_Sales\_skewness1,

(select

round((sum(pow(RtnMPR - (select avg(RtnMPR) from new\_table),3))/(count(\*) \*

pow(stddev(RtnMPR),3))),2) as RtnMPR\_skewness from new\_table)

RtnMPR\_skewness1;



**4. Calculating the fourth moment (kurtosis) for the dataset.**

**Kurtosis:**

select

kurtosis\_quantity1.kurtosis\_quantity,

kurtosis\_Returnquantity1.kurtosis\_Returnquantity,

kurtosis\_Final\_cost1.kurtosis\_Final\_cost,

kurtosis\_Final\_Sales1.kurtosis\_Final\_Sales,

kurtosis\_RtnMPR1.kurtosis\_RtnMPR

from

(SELECT

ROUND((SUM(POWER(Quantity - (SELECT AVG(Quantity) FROM new\_table), 4)) / (COUNT(Quantity) \* POWER(STDDEV(Quantity),

4))), 2) AS kurtosis\_quantity

from new\_table) kurtosis\_quantity1,

(SELECT

ROUND((SUM(POWER(ReturnQuantity - (SELECT AVG(ReturnQuantity) FROM new\_table), 4)) / (COUNT(ReturnQuantity) \* POWER(STDDEV(ReturnQuantity),

4))), 2) AS kurtosis\_Returnquantity

from new\_table) kurtosis\_Returnquantity1,

(SELECT

ROUND((SUM(POWER(Final\_cost - (SELECT AVG(Final\_cost) FROM new\_table), 4)) / (COUNT(Final\_cost) \* POWER(STDDEV(Final\_cost),

4))), 2) AS kurtosis\_Final\_cost

from new\_table) kurtosis\_Final\_cost1,

(SELECT

ROUND((SUM(POWER(Final\_Sales - (SELECT AVG(Final\_Sales) FROM new\_table), 4)) / (COUNT(Final\_Sales) \* POWER(STDDEV(Final\_Sales),

4))), 2) AS kurtosis\_Final\_Sales

from new\_table) kurtosis\_Final\_Sales1,

(SELECT

ROUND((SUM(POWER(RtnMPR - (SELECT AVG(RtnMPR) FROM new\_table), 4)) / (COUNT(RtnMPR) \* POWER(STDDEV(RtnMPR),

4))), 2) AS kurtosis\_RtnMPR

from new\_table) kurtosis\_RtnMPR1;



**Comparison table showing the business decisions results for the unclean and clean data:**

|  |  |  |
| --- | --- | --- |
| **Parameter** | **Uncleaned Data** | **Cleaned\_Data** |
| **Mean** | **Quantity -2.25** | **Quantity-1.78** |
|  | **ReturnQuantity -0.30** | **ReturnQuantity -0.19** |
|  | **Final\_cost-124.43** | **Final\_cost-88.47** |
|  | **Final\_Sales-232.79** | **Final\_Sales-165.01** |
|  | **RtnMPR-29.95** | **RtnMPR-13.85** |
| **Median** | **Quantity-3.00** | **Quantity-1.00** |
|  | **ReturnQuantity -0.00** | **ReturnQuantity -0.00** |
|  | **Final\_cost-141** | **Final\_cost-52.99** |
|  | **Final\_Sales-53.57** | **Final\_Sales-56.44** |
|  | **RtnMPR-0-00** | **RtnMPR-0.00** |
| **Mode** | **Quantity-1** | **Quantity-1** |
|  | **ReturnQuantity -0** | **ReturnQuantity -0** |
|  | **Final\_cost-49.352** | **Final\_cost-49.352** |
|  | **Final\_Sales-60.8** | **Final\_Sales-60.8** |
|  | **RtnMPR-0** | **RtnMPR-0** |
| **Variance** | **Quantity-27.14** | **Quantity-3.32** |
|  | **ReturnQuantity 2.17** | **ReturnQuantity -0.41** |
|  | **Final\_cost-221685.41** | **Final\_cost-17736.69** |
|  | **Final\_Sales-458223.23** | **Final\_Sales-71134.19** |
|  | **RtnMPR-34178.39** | **RtnMPR-2596.22** |
| **StandardDeviation** | **Quantity-5.21** | **Quantity-1.82** |
|  | **ReturnQuantity -1.67** | **ReturnQuantity -0.64** |
|  | **Final\_cost-470.83** | **Final\_cost-133.18** |
|  | **Final\_Sales-676.93** | **Final\_Sales-266.71** |
|  | **RtnMPR-184.87** | **RtnMPR-49.96** |
| **Range** | **Quantity-150** | **Quantity-17** |
|  | **ReturnQuantity -50** | **ReturnQuantity -5** |
|  | **Final\_cost-33138** | **Final\_cost-1363.11** |
|  | **Final\_Sales-39490** | **Final\_Sales-2143.64** |
|  | **RtnMPR-8014** | **RtnMPR-568.18** |
| **Skewness** | **Quantity-11.27** | **Quantity-2.91** |
|  | **ReturnQuantity -16.88** | **ReturnQuantity -4.23** |
|  | **Final\_cost-34.22** | **Final\_cost-5.86** |
|  | **Final\_Sales-21.09** | **Final\_Sales-4.3** |
|  | **RtnMPR-15.58** | **RtnMPR5.11** |
| **Kurtosis** | **Quantity-180.5**   |  |  |  |  |  |  | | --- | --- | --- | --- | --- | --- | |  |  |  |  |  |  | | **Quantity-14.7** |
|  | **ReturnQuantity -398.03** | **ReturnQuantity -23.43** |
|  | **Final\_cost-1984.44** | **Final\_cost-41.6** |
|  | **Final\_Sales-947.71** | **Final\_Sales-24.76** |
|  | **RtnMPR-395.28** | **RtnMPR-35.42** |

**Observation: Overall, the results indicate that the unclean data exhibits higher mean, variance, standard deviation, range, skewness, and kurtosis values compared to the clean data. This suggests greater inconsistencies, variability, and potential outliers in the unclean data. Cleaning the data has resulted in more stable and normalized distributions with reduced variability and potential biases, making it more reliable for business decision-making.**