Assignment – 2 (1P21CS006)

1. Write a map only algorithm which will read the original dataset as input and filter out all the records which have event\_epoch\_time, user\_id, device\_id, user\_agent as NULL.

**Solution:**

LS = split(value,’\t’)

event\_epoch\_time = LS[1]

user\_id = LS[2]

device\_id = LS[3]

user\_agent = LS[4]

if(event\_epoch\_time != NULL)

write(event\_epoch\_time)

if(user\_id != NULL)

write(user\_id)

if(device\_id != NULL)

write(device\_id)

if(user\_agent != NULL)

write(user\_agent)

2. An algorithm to read the user agent and extract OS Version and Platform from it.

**Solution:**

LS = split(user\_agent, ’ : ’)

Platform = LS[1]

OS\_version = LS[2]

write(platform, OS\_version)

3. Assume there is a predefined method named getCounter(String name) which takes a name as the parameter and creates a global counter variable of the same name if already not created. This global counter variable is accessible to all the map tasks. To increment the value of a counter the method to be used is incrementBy(integer num). Here “num” is the number by which we want to increment the global variable. So the syntax to increment the value of a counter is: getCounter(“Orders”).incrementBy(1)

Using the above info write algorithms to perform below-mentioned tasks:

1. Find out the number of veg and non-veg pizzas sold

**Solution:**

LS = split(value,’\t’)

Veg = LS[12]

Order = LS[11]

If(Veg == “Y” & Order == “Delivered”)

getCounter(“isVeg”).incrementBy(1)

If(Veg == “N” & Order == “Delivered”)

getCounter(“notVeg”).incrementBy(1)

write(“Veg: ” isVeg”, Non - Veg: ” notVeg)

2. Find out the size wise distribution of pizzas sold

**Solution:**

LS = split(value,’\t’)

Size = LS[7]

Order = LS[11]

If(Size == “M” & Order == “Delivered”)

getCounter(“mid”).incrementBy(1)

If(Size == “R” & Order == “Delivered”)

getCounter(“reg”).incrementBy(1)

If(Size == “L” & Order == “Delivered”)

getCounter(“lar”).incrementBy(1)

write(“medium: “mid, ”Regular: ”reg, ”Large: “lar)

3. Find out how many cheese burst pizzas were sold

**Solution:**

LS = split(value,’\t’)

Cheese\_burst = LS[6]

Order = LS[11]

If(Cheese\_burst == “Y” & Order == “Delivered”)

getCounter(“having\_cheese”).incrementBy(1)

If(Cheese\_burst == “N” & Order == “Delivered”)

getCounter(“not having\_cheese”).incrementBy(1)

write(“Cheese\_burst: ” having\_cheese, “Non Cheese\_burst: ” not having\_cheese)

4. Find out how many small cheese burst pizzas were sold. Ideally, the count should be 0

**Solution:**

LS = split(value,’\t’)

Cheese\_burst = LS[6]

Order = LS[11]

Size = LS[7]

If(Cheese\_burst == “Y” & Size == “S” & Order == “Delivered”)

getCounter(“small\_cheese\_burst”).incrementBy(1)

write(“small cheese burst: ” small\_cheese\_burst)

because cheese burst is available for medium and large

5. Find out the number of cheese burst pizzas whose cost is below Rs 500

**Solution:**

LS = split(value,’\t’)

Price = LS[9]

If(Price < 500)

getCounter(“price\_below\_500”).incrementBy(1)

write(price\_below\_500)

4. Assume that the predefined method getCounter does not exist. Write the updated algorithms for the

tasks in point-3.

**All 5 questions Solution:**

LS = split(value, ’\t’)

isCheeseBurst = LS[6]

Size = LS[7]

Price = LS[9]

Order\_Event = LS[11]

isVeg = LS[12]

1.

If(isVeg == “Y” & Order\_Event == “Delivered”)

Yes\_cnt = 0

For i in valueList.length

Yes\_cnt+=1

Write(Yes\_cnt)

If(isVeg == “N” & Order\_Event == “Delivered”)

No\_cnt = 0

For i in valueList.length

No\_cnt+=1

Write(No\_cnt)

2.

If(Size == “M” & Order\_Event == “Delivered”)

mid\_cnt = 0

For i in valueList.length

Mid\_cnt+=1

Write(mid\_cnt)

If(Size == “R” & Order\_Event == “Delivered”)

reg\_cnt = 0

For i in valueList.length

reg\_cnt+=1

Write(reg\_cnt)

If(Size == “L” & Order\_Event == “Delivered”)

lar\_cnt = 0

For i in valueList.length

lar\_cnt+=1

Write(lar\_cnt)

3.

If(isCheeseBurst == “Y” & Order\_Event == “Delivered”)

Yes\_cnt = 0

For i in valueList.length

Yes\_cnt+=1

Write(Yes\_cnt)

If(isCheeseBurst == “N” & Order\_Event == “Delivered”)

No\_cnt = 0

For i in valueList.length

No\_cnt+=1

Write(No\_cnt)

4.

If(Cheese\_burst == “Y” & Size == “S” & Order == “Delivered”)

cnt = 0

For i in valueList.length

cnt+=1

Write(cnt)

5.

If(Price < 500)

cnt = 0

For i in valueList.length

cnt+=1

Write(cnt)