**Q-1) How many different types of organization groups without duplicate?**

**display(**

**df .select("Organization Group")**

**.distinct()**

**.orderBy("Organization Group"))**

**Q-2) How many different types of Department group without duplicate?**

**display(**

**df .select("Department")**

**.distinct()**

**.orderBy("Department "))**

**Q-3) Show sum of overtime based on Department group?**

**display(**

**df**

**.select(df["Department"],df["Overtime"])**

**.groupBy("Department")**

**.sum() )**

**Q-4)** Calculate **Final salary considering salary \* overtime**

**display(**

**df .select((df["Salaries"]+df["Overtime"]).alias("FinalSalary")))**

**Q-5) Show sum of overtime based on Department group?**

**display(**

**df**

**.select(df["Department"],df["Overtime"])**

**.groupBy("Department")**

**.sum("Overtime")**

**)**

**Q-6) Show me the list of overtime 0?**

**Solution 1 :**

**df.filter(df["Overtime"]==0).show()**

**Solution 2 :**

**display(**

**df .select(df["Job"])**

**.filter(df["Overtime"]==0)**

**)**

**Q-7) Show me the total cost of health/dental?**

**display(**

**df**

**.select(df["Health/Dental"])**

**.groupBy("Health/Dental")**

**.sum("Health/Dental")**

**)**

**Q-8) Show me the total compensation according to organization group?**

**display(**

**df**

**.select(df["Organization Group"],df["Total Compensation"])**

**.groupBy("Organization Group")**

**.sum("Total Compensation")**

**)**

**Q-9) Show me the salary who has more than 70000?**

**display(**

**df**

**.select(df["Job"],df["Salaries"])**

**.filter(df["Salaries"] >70000)**

**)**

**Q-10) Show the count of question 9?**

**display(**

**df**

**.groupBy(df["Salaries"] >70000).count()**

**)**