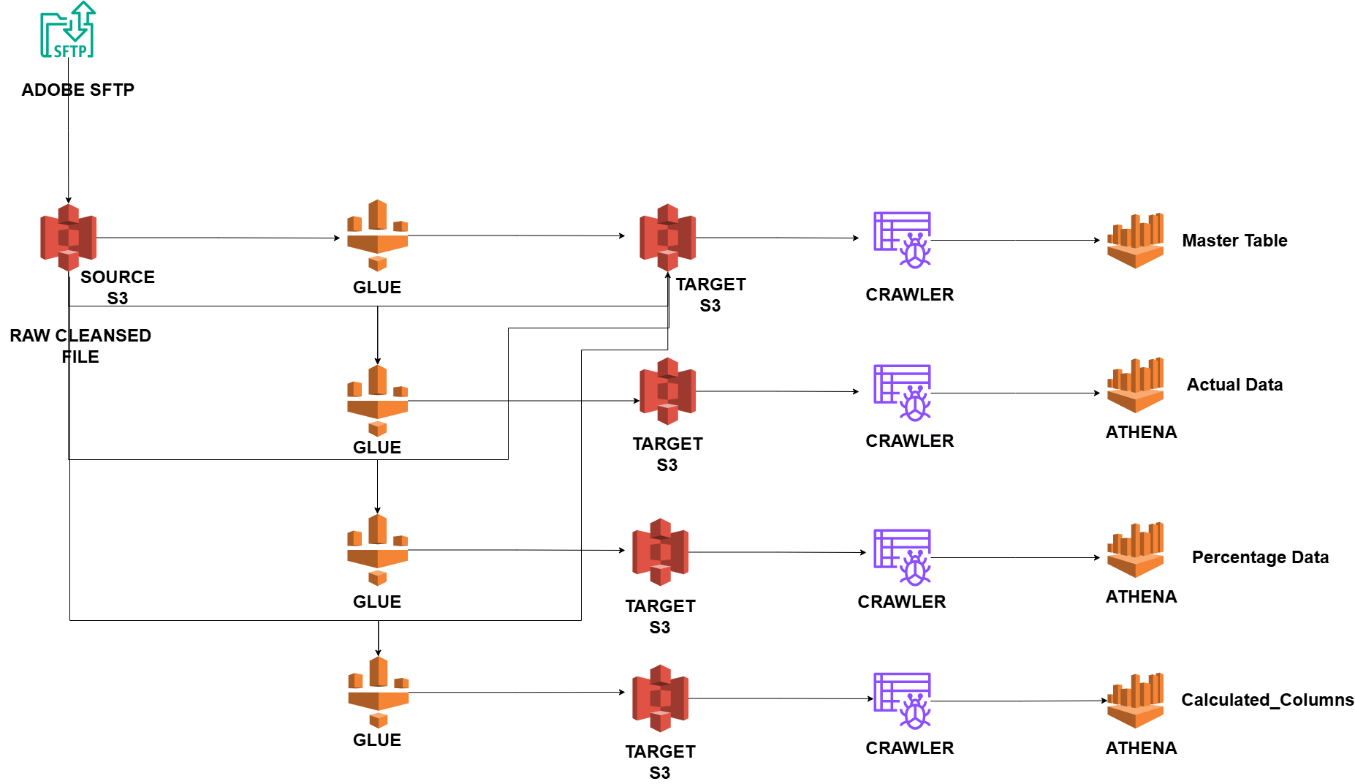
**ARCHITECTURE DIAGRAM FOR EXISITING HEALTH DASHBOARD BACKEND PROCESS:**



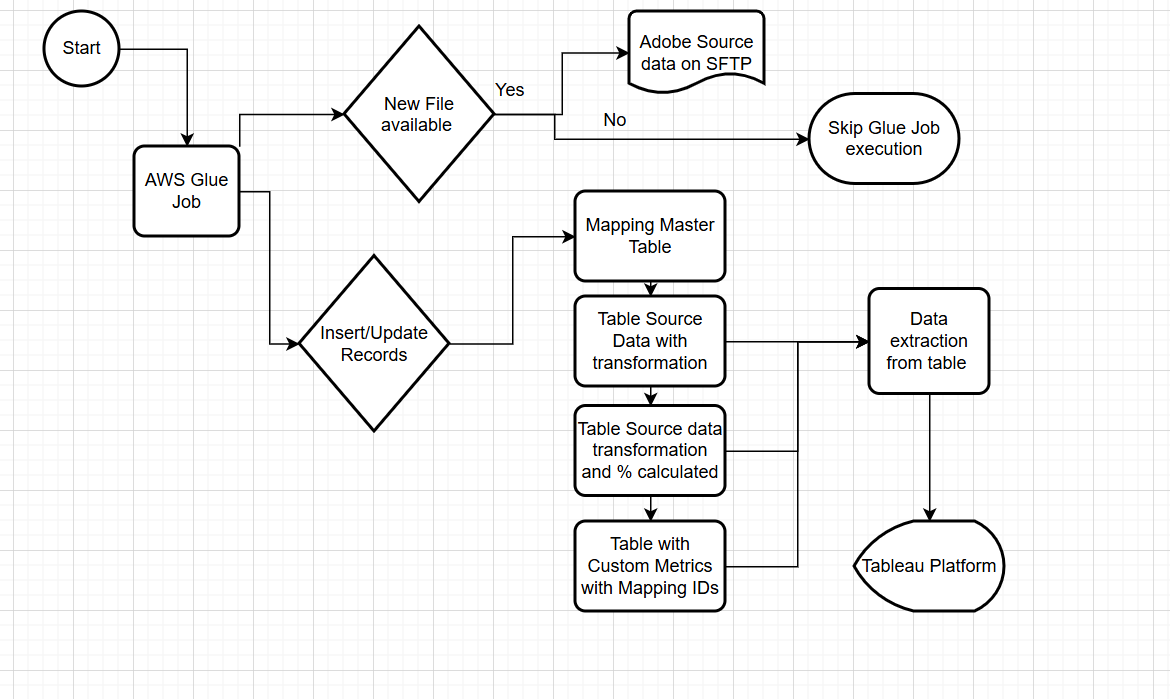
**DESCRIPTION:**

* Reading a **raw cleansed csv file** and generated unique\_id for the combination of metric + OS and load it as Master table **smarthelp\_master** using **HC\_Mapping\_Master** this glue job
* Reading a raw cleansed csv file & **HC\_Mapping\_Master** Athena table, pivoted the Date columns as Row format and the actual numbers as Values.
* Reading a raw cleansed csv file & **HC\_Mapping\_Master** Athena table, calculating percentage and Both Metrics and Pivoted the Date columns as Ror Format and Percentages as Values.
* Reading a raw cleansed csv file & **HC\_Mapping\_Master** Athena table, calculating the required columns like 1st Quartile, 3rd Quartile, IQR, Upper, Lower, Yesterday, Last 7 Days, Last 30 Days, Last 7 Days%, Last 30 Days %.

**QUESTIONS ABOUT HOW SOURCE FILES REFRESHING:**

* How the files are refreshing in on Adobe Source in SFTP Daily/Weekly/Monthly?
* Does the data refreshing in same file / adding new file with new data in Adobe Source in SFTP?
* Does COX team itself loading the latest data into s3 from Adobe Source SFPT or we need to develop that functionality?

**DATA FLOW DIAGRAM FOR EXISITING HEALTHSCORE DASHBOARD BACKEND PROCESS :**



**EXISITNG EXCEL HEALTHSCORE DASHBOARD : **

**CLEANSED INPUT SOURCE FILE FOR SMART HELP : **

**ATHENA TABLES :**

* smarthelp\_master
* smarthelp\_data\_actual
* smarthelp\_data\_calc
* smarthelp\_data\_metric

**TABLE STRUCTURE:**

|  |  |  |
| --- | --- | --- |
| **Table Name - HC\_Mapping\_Master** | | |
| **Field Name** | **Data type** | **Comment** |
| ID | Int | Auto generated number |
| OS Type | Text |  |
| Metrics | Text |  |
| Sum Metric | Text |  |
| Feature | Text |  |
| Metrics Name | Text | This could be a different values |
| Display Name | Text | This values should be shown to Dashboard. Also display name may be different for OS Type, Metrics or Sub Metrics |

|  |  |  |
| --- | --- | --- |
| **Table Name - HC\_Source\_Data\_Actual** | | |
| **Field Name** | **Data type** | **Comment** |
| ID | Int | Taken from Master table for the OS Type and Metrics name |
| Date | Date | Date transfomred from Column to Row |
| Value | Int | Actual Value saved for respective date and OS Type and Metrics Name. We will have ID only from Master table |

|  |  |  |
| --- | --- | --- |
| **Table Name - HC\_Source\_Data\_Calc** | | |
| **Field Name** | **Data type** | **Comment** |
| ID | Int | Taken from Master table for the OS Type and Metrics name |
| Date | Date | Date transfomred from Column to Row |
| Value | Int | % Value saved for respective date, calcualted by python code and OS Type and Metrics Name. We will have ID only from Master table |

|  |  |  |
| --- | --- | --- |
| **Table Name - HC\_DB\_Metrics\_Cal** | | |
| **Field Name** | **Data type** | **Comment** |
| ID | Int | Taken from Master table for the OS Type and Metrics name |
| Date | Date | Date transfomred from Column to Row |
| Upper limit |  | I am not too sure about the column name but we should add all the column remain from python code output at the last after saving date and values ( actual and %) |
| Lower Limit |  |
| 7Days |  |
| 30Days |  |

**GLUE JOB FUCNTIONALITIES EXISTING APPROACH :**

**HC\_Mapping\_Master job : **

* Reading **Cleansed Input file** with only Smart\_Help Feature
* Generating a sequential **ID** column for each Metrics
* Adding **sub\_metrics, display\_name, feature, sub\_display, new\_metrics\_name** manually hardcoded the values in the code
* Generating a **create\_dt** column with current\_date function
* Loading those columns in s3 path and created Athena tables using crawler

**HC\_Source\_Data\_Actual : **

* Reading **Cleansed Input file** with only Smart\_Help Feature & Reading **HC\_Mapping\_Master\_table** data and join both the tables
* Pivoting the **Date columns** into **row format** and one more column **Values** to store the **actual number** from cleansed input file
* Loading only **ID, create\_dt, Date, Value** these columns in another s3 path & created Athena tables using crawler.

**HC\_Source\_Data\_Calc Job : **

* Reading **Cleansed Input file** with only Smart\_Help Feature & Reading **HC\_Mapping\_Master\_table** data and join both the tables
* Calculating the **Percentage** for each **Date Columns** for all the metrics based on the respective **Numerator/ Denominator** with hardcoded metric names

**Logic :** df\_pivot['1st common reason (Unknown) %'] = (df\_pivot['1st common reason (Unknown)'] / df\_pivot['Saw a Result'] \* 100).round(3)

* Calculating **Both Metrics(Android & iOS)** rows for each metrics
* Pivoting the **Date columns** into **row format** and one more column **Values** to store the **Percenatage values** from calculations.

**HC\_ job\_4new-copy :**

* Generating the calculated columns **like ID, 1st Quartile, 3rd Quartile, IQR, Upper, Lower, Yesterday, Last 7 days, Last 30 days, % Last 7 Days, % Last 30 Days, Date, Value, create\_dt**

These excel files has logic for all the calculated columns

****

**GLUE JOB FUCNTIONALITIES DYNAMIC APPROACH**

**Python Script : CLEANSED INPUT FILE ALL FEATURE : **

**First Table :  Second Table: **

* Reading cleansed input file with all the features
* Generating **metric\_id, display\_names, feature\_name, unique\_identifier, level1,levele2,level3,level4, hierarchy\_id, parent\_id\_old, metric\_sequence\_num,levelno, feature\_id, parent\_id, create\_dt**
* These columns data are loaded into new s3 path and create Athena table using crawler **health\_score\_all\_metrics\_master\_table**
* From the **cleansed input file**, pivoting the **Date columns** **into row format** with the **actual** **numbers** stored in **Values** column.
* Based on the **hierarchy\_id, parent\_id\_old,** numerators & denominators can be determined and the percentage calculation of each metrics is calculated.

**Logic** : **hierarchy\_id** rows or index of every metrics is **Numerator**

In the same index or row, the **parent\_id\_old** values has to check in **hierarchy\_id,** if the values present that index or row is the **Denominator**

Calculating Both Metrics & 4th table columns (1st Quartile, 3rd Quartile, IQR, Upper, Lower, Yesterday, Last 7 days, Last 30 days, % Last 7 Days, % Last 30 Days) – Not yet Started

**for Devices Page Visits** = Summetion of Devices Page Visits Apple iOS and Google Android ,

**for Started Troubleshooting (Solution Center) =**

(Started Troubleshooting (Solution Center) %, Operating System Type= Apple iOS \*

Devices Page Visits Operating System Type =Apple iOS

+

(Started Troubleshooting (Solution Center) %, Operating System Type Google =Android \*

Devices Page Visits Operating System Type = Google Android)

/

Devices Page Visits ,Operating System = Type Apple iOS Addition with

Devices Page Visits, Operating System = Type Google Android)

**for Saw a Result =**

(Saw a Result, Operating System Type Apple iOS Addition with

Saw a Result Operating System Type google Android)

Division

(Started Troubleshooting (Solution Center) ,Operating System Type= Apple iOS Addition with

Started Troubleshooting (Solution Center), Operating System Type= Google Android)

**For 1st common reason (Unknown) =**

(1st common reason (Unknown), Operating System Type Apple iOS Addition with

1st common reason (Unknown) Operating System Type google Android)

Division

(Saw a Result ,Operating System Type= Apple iOS Addition with

Saw a Result, Operating System Type= Google Android)

**2nd common reason (Disconnected) =**

(2nd common reason (Disconnected), Operating System Type Apple iOS Addition with

2nd common reason (Disconnected) Operating System Type google Android)

Division

(Saw a Result ,Operating System Type= Apple iOS Addition with

Saw a Result, Operating System Type= Google Android)

**3rd common reason (Incomplete Install) =**

(3rd common reason (Incomplete Install), Operating System Type Apple iOS Addition with

3rd common reason (Incomplete Install) Operating System Type google Android)

Division

(Saw a Result ,Operating System Type= Apple iOS Addition with

Saw a Result, Operating System Type= Google Android)

**Did a Device Reboot =**

(Did a Device Reboot %, Operating System Type= Apple iOS \*

Devices Page Visits Operating System Type =Apple iOS

+

(Did a Device Reboot %, Operating System Type= Google Android \*

Devices Page Visits Operating System Type = Google Android)

/

Devices Page Visits ,Operating System = Type Apple iOS Addition with

Devices Page Visits, Operating System = Type Google Android)

**Started Proactive and Preventive (P&P) =**

(Started Proactive and Preventive (P&P) %, Operating System Type= Apple iOS \*

Devices Page Visits Operating System Type =Apple iOS

+

(Started Proactive and Preventive (P&P) %, Operating System Type= Google Android \*

Devices Page Visits Operating System Type = Google Android)

/

Devices Page Visits ,Operating System Type= Apple iOS Addition with

Devices Page Visits, Operating System Type= Google Android)

**Started P&P troubleshooting =**

(Started P&P troubleshooting, Operating System Type= Apple iOS Addition with

Started P&P troubleshooting Operating System Type= Google Android)

Division

(Started Proactive and Preventive (P&P) ,Operating System Type= Apple iOS Addition with

Started Proactive and Preventive (P&P), Operating System Type= Google Android)

**Reset Complete =**

(Reset Complete, Operating System Type= Apple iOS Addition with

Reset Complete, Operating System Type= Google Android)

Division

(Started Proactive and Preventive (P&P) ,Operating System Type= Apple iOS Addition with

Started Proactive and Preventive (P&P), Operating System Type= Google Android)

**Schedule Tech =**

(Schedule Tech, Operating System Type= Apple iOS Addition with

Schedule Tech, Operating System Type= Google Android)

Division

(Started Proactive and Preventive (P&P) ,Operating System Type= Apple iOS Addition with

Started Proactive and Preventive (P&P), Operating System Type= Google Android)

**Connect with Agent =**

(Connect with Agent, Operating System Type= Apple iOS Addition with

Connect with Agent, Operating System Type= Google Android)

Division

(Started Proactive and Preventive (P&P) ,Operating System Type= Apple iOS Addition with

Started Proactive and Preventive (P&P), Operating System Type= Google Android)

**Interacted with Chat**

(Interacted with Chat, Operating System Type= Apple iOS Addition with

Interacted with Chat, Operating System Type= Google Android)

Division

(Started Proactive and Preventive (P&P) ,Operating System Type= Apple iOS Addition with

Started Proactive and Preventive (P&P), Operating System Type= Google Android)

**Clicked Billing & Account Support =**

(Clicked Billing & Account Support %, Operating System Type= Apple iOS \*

Devices Page Visits Operating System Type =Apple iOS

+

(Clicked Billing & Account Support %, Operating System Type= Google Android \*

Devices Page Visits Operating System Type = Google Android)

/

Devices Page Visits ,Operating System Type= Apple iOS Addition with

Devices Page Visits, Operating System Type= Google Android)

**Clicked Internet Support =**

(Clicked Internet Support %, Operating System Type= Apple iOS \*

Devices Page Visits Operating System Type =Apple iOS

+

(Clicked Internet Support %, Operating System Type= Google Android \*

Devices Page Visits Operating System Type = Google Android)

/

Devices Page Visits ,Operating System Type= Apple iOS Addition with

Devices Page Visits, Operating System Type= Google Android)

**Clicked TV Support =**

(Clicked TV Support %, Operating System Type= Apple iOS \*

Devices Page Visits Operating System Type =Apple iOS

+

(Clicked TV Support %, Operating System Type= Google Android \*

Devices Page Visits Operating System Type = Google Android)

/

Devices Page Visits ,Operating System Type= Apple iOS Addition with

Devices Page Visits, Operating System Type= Google Android)

**Clicked Homelife Support =**

(Clicked Homelife Support %, Operating System Type= Apple iOS \*

Devices Page Visits Operating System Type =Apple iOS

+

(Clicked Homelife Support %, Operating System Type= Google Android \*

Devices Page Visits Operating System Type = Google Android)

/

Devices Page Visits ,Operating System Type= Apple iOS Addition with

Devices Page Visits, Operating System Type= Google Android)

**Interacted with Chat\_1 =**

(Interacted with Chat\_1 %, Operating System Type= Apple iOS \*

Devices Page Visits Operating System Type =Apple iOS

+

(Interacted with Chat\_1 %, Operating System Type= Google Android \*

Devices Page Visits Operating System Type = Google Android)

/

Devices Page Visits ,Operating System Type= Apple iOS Addition with

Devices Page Visits, Operating System Type= Google Android)