{

    "name": "Dataflow Summary",

    "properties": {

        "type": "MappingDataFlow",

        "typeProperties": {

            "sources": [

                {

                    "dataset": {

                        "referenceName": "ds\_SQL\_2020",

                        "type": "DatasetReference"

                    },

                    "name": "source1"

                },

                {

                    "dataset": {

                        "referenceName": "ds\_SQL\_2021",

                        "type": "DatasetReference"

                    },

                    "name": "source2"

                }

            ],

            "sinks": [

                {

                    "dataset": {

                        "referenceName": "ds\_SQL\_Summary",

                        "type": "DatasetReference"

                    },

                    "name": "sink1"

                },

                {

                    "dataset": {

                        "referenceName": "ds\_sumarry",

                        "type": "DatasetReference"

                    },

                    "name": "sink2"

                }

            ],

            "transformations": [

                {

                    "name": "union1"

                },

                {

                    "name": "filter1"

                },

                {

                    "name": "derivedColumn1"

                },

                {

                    "name": "aggregate1"

                }

            ],

            "scriptLines": [

                "parameters{",

                "     dataflow\_param\_fiscalyear as integer (2020)",

                "}",

                "source(output(",

                "          FiscalYear as integer,",

                "          PayrollNumber as integer,",

                "          AgencyID as string,",

                "          AgencyName as string,",

                "          EmployeeID as string,",

                "          LastName as string,",

                "          FirstName as string,",

                "          AgencyStartDate as date,",

                "          WorkLocationBorough as string,",

                "          TitleCode as string,",

                "          TitleDescription as string,",

                "          LeaveStatusasofJune30 as string,",

                "          BaseSalary as double,",

                "          PayBasis as string,",

                "          RegularHours as double,",

                "          RegularGrossPaid as double,",

                "          OTHours as double,",

                "          TotalOTPaid as double,",

                "          TotalOtherPay as double",

                "     ),",

                "     allowSchemaDrift: true,",

                "     validateSchema: false,",

                "     isolationLevel: 'READ\_UNCOMMITTED',",

                "     format: 'table') ~> source1",

                "source(output(",

                "          FiscalYear as integer,",

                "          PayrollNumber as integer,",

                "          AgencyCode as string,",

                "          AgencyName as string,",

                "          EmployeeID as string,",

                "          LastName as string,",

                "          FirstName as string,",

                "          AgencyStartDate as date,",

                "          WorkLocationBorough as string,",

                "          TitleCode as string,",

                "          TitleDescription as string,",

                "          LeaveStatusasofJune30 as string,",

                "          BaseSalary as double,",

                "          PayBasis as string,",

                "          RegularHours as double,",

                "          RegularGrossPaid as double,",

                "          OTHours as double,",

                "          TotalOTPaid as double,",

                "          TotalOtherPay as double",

                "     ),",

                "     allowSchemaDrift: true,",

                "     validateSchema: false,",

                "     isolationLevel: 'READ\_UNCOMMITTED',",

                "     format: 'table') ~> source2",

                "source1, source2 union(byName: true)~> union1",

                "union1 filter(toInteger(FiscalYear) >= $dataflow\_param\_fiscalyear) ~> filter1",

                "filter1 derive(TotalPaid = RegularGrossPaid + TotalOTPaid+TotalOtherPay) ~> derivedColumn1",

                "derivedColumn1 aggregate(groupBy(AgencyName,",

                "          FiscalYear),",

                "     TotalPaid = sum(TotalPaid)) ~> aggregate1",

                "aggregate1 sink(allowSchemaDrift: true,",

                "     validateSchema: false,",

                "     input(",

                "          FiscalYear as integer,",

                "          AgencyName as string,",

                "          TotalPaid as double",

                "     ),",

                "     deletable:false,",

                "     insertable:true,",

                "     updateable:false,",

                "     upsertable:false,",

                "     truncate:true,",

                "     format: 'table',",

                "     skipDuplicateMapInputs: true,",

                "     skipDuplicateMapOutputs: true,",

                "     errorHandlingOption: 'stopOnFirstError') ~> sink1",

                "aggregate1 sink(allowSchemaDrift: true,",

                "     validateSchema: false,",

                "     truncate: true,",

                "     umask: 0022,",

                "     preCommands: [],",

                "     postCommands: [],",

                "     skipDuplicateMapInputs: true,",

                "     skipDuplicateMapOutputs: true,",

                "     partitionBy('hash', 1)) ~> sink2"

            ]

        }

    }

}