

Name: Shashwat Karn

19BCE2625

VIT VELLORE,

Smarinternzzz

Assignment – 03

Diabetes Dataset

New project

Define details

Name

Project name

Description

Project description

Choose project options

☐ Restrict who can be a collaborator ⓘ

Project includes integration with [Cloud Object Storage](#) for storing project assets.

Storage

Cloud Object Storage-hv

Cancel

Create

Create New Project


Projects / Classification of Diabetes

Launch IDE

OverviewAssetsJobsManage

Assets

No assets created with tools yet.



[View all](#)

Resource usage ⓘ

For this month in this project

0 CPU

Project history

No notifications ⓘ

You will see your most recent notifications here.

Readme ⓘ

Type project notes, reminders, or instructions

Data in this project

Drop data files here or browse for files to upload

diabetes.csv

Uploaded!

Dismiss

The Dataset

The screenshot shows the 'Assets' tab in a project interface. The breadcrumb path is 'Projects / Classification of Diabetes'. The top navigation bar includes 'Overview', 'Assets', 'Jobs', and 'Manage'. The 'Assets' tab is active, displaying a search bar and buttons for 'Add asset' and 'New asset'. On the left, a sidebar shows '1 assets' and 'Asset types' with 'Data' selected. The main area, titled 'Data asset', contains a table with one item: 'diabetes.csv' (CSV format), last modified 'Now' by 'Aditya Yadav (You)'. A right-hand panel titled 'Data in this project' has a dashed box for dropping files. At the bottom, it shows 'Items per page: 20' and '1~1 of 1 items'.

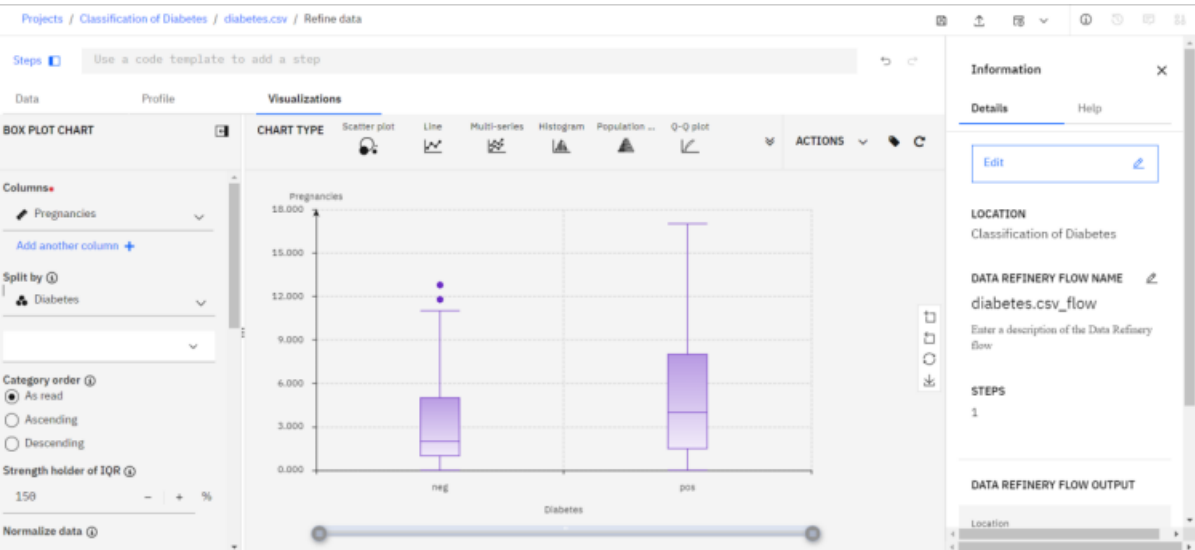
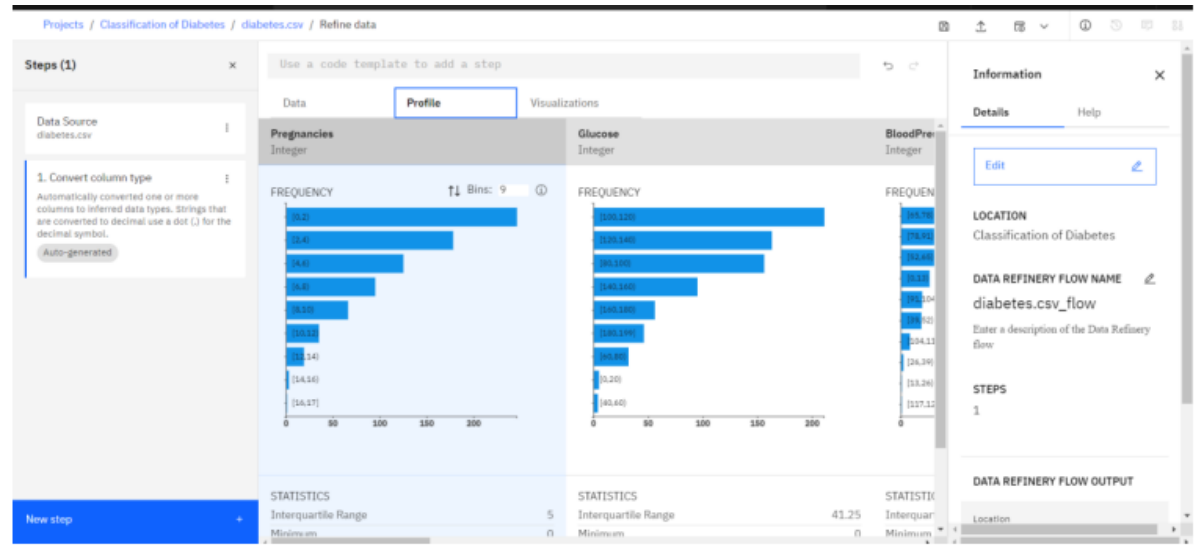
Name	Last modified
diabetes.csv CSV	Now Aditya Yadav (You)

Refine Data

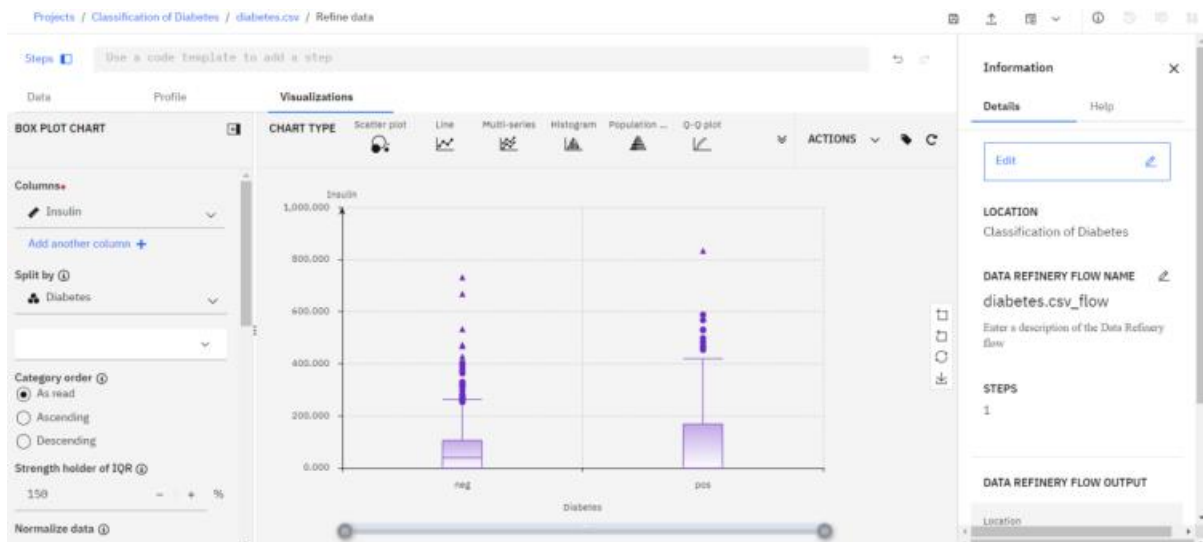
The screenshot shows the 'Refine data' interface for the 'diabetes.csv' asset. The breadcrumb path is 'Projects / Classification of Diabetes / diabetes.csv / Refine data'. The top navigation bar includes 'Overview', 'Assets', 'Jobs', and 'Manage'. The 'Refine data' tab is active, displaying a search bar and buttons for 'Add asset' and 'New asset'. On the left, a sidebar shows 'Steps (1)' with a step '1. Convert column type' (Auto-generated). The main area, titled 'Data', contains a table with 14 rows of data. A right-hand panel titled 'Information' shows details for the 'diabetes.csv_flow' data refinery flow, including its location and steps. At the bottom, it shows 'SOURCE FILE: diabetes.csv' and 'FULL DATA SET: 768 rows'.

	Pregnancies	Glucose	BloodPress...	SkinThickn...	Insulin
	Integer	Integer	Integer	Integer	Integer
1	6	148	72	35	0
2	1	85	66	29	0
3	8	183	64	0	0
4	1	89	66	23	94
5	0	137	40	35	168
6	5	116	74	0	0
7	3	78	50	32	88
8	10	115	0	0	0
9	2	197	70	45	543
10	8	125	96	0	0
11	4	110	92	0	0
12	10	168	74	0	0
13	10	139	80	0	0
14	1	189	60	23	846

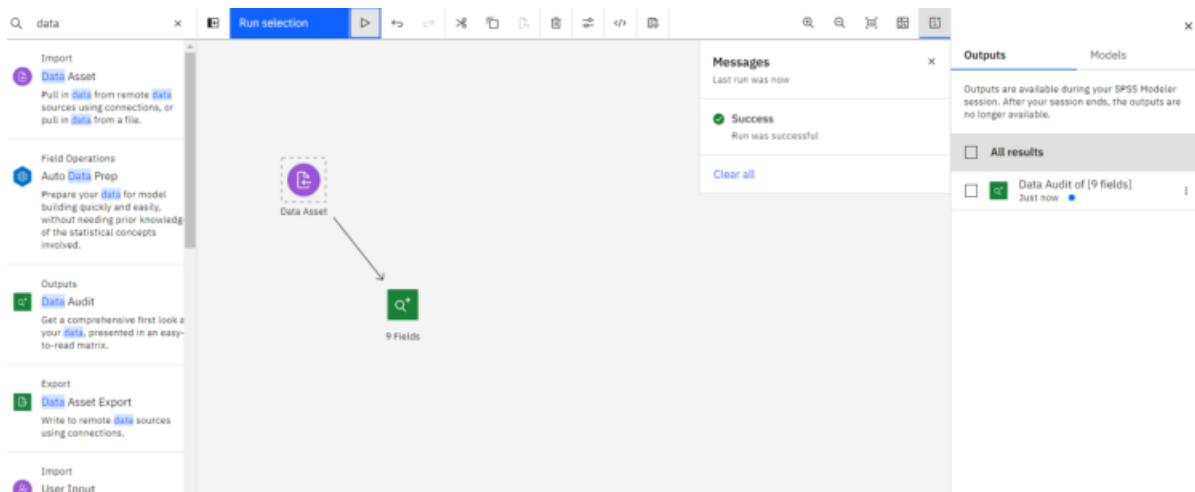
Profile of our data



Graph of Insulin and Database



Create SSP modeller and add dataset and audit










Output of Audit

Projects / Classification of Diabetes / Classification Of Diabetes

View Output: Data Audit of [9 fields]

Compare

	Field	Graph	Measurement	Min	Max	Mean	Std. Dev	Skewness	Unique	Valid
1	Pregnancies		Continuous	0	17	3.845	3.370	0.902	--	768
2	Glucose		Continuous	0	199	120.895	31.973	0.174	--	768
3	BloodPressure		Continuous	0	122	69.105	19.356	-1.844	--	768
4	SkinThickness		Continuous	0	99	20.536	15.952	0.109	--	768
5	Insulin		Continuous	0	846	79.799	115.244	2.272	--	768
6	BMI		Continuous	0.000	67.100	31.993	7.884	-0.429	--	768
7	Diabetes		Continuous	0.000	1.000	0.350	0.481	0.555	--	768

Target

Projects / Classification of Diabetes / Classification Of Diabetes

type

Field Operations

- Type: Specify field metadata and properties that are invaluable to modeling.
- Filter: Replace field values and change storage. Often used with a Type node to replace missing values.
- Multiplot: A special type of plot that displays multiple Y fields over a single X field.
- Distribution: Shows the occurrence of symbolic (non-numeric) values, such as mortgage type or gender, in a dataset.

Data Asset

Type

9 Fields

Type

Settings

Read values Clear values

Find in column Field

Field	Measure	Role	Value mode	Values
<input type="checkbox"/> # Pregnancies	Continuous	Input	Read	
<input type="checkbox"/> # Glucose	Continuous	Input	Read	
<input type="checkbox"/> # BloodPressure	Continuous	Input	Read	
<input type="checkbox"/> # SkinThickness	Continuous	Input	Read	
<input type="checkbox"/> # Insulin	Continuous	Input	Read	
<input type="checkbox"/> BMI	Continuous	Input	Read	
<input type="checkbox"/> DiabetesPedigree	Continuous	Input	Read	
<input type="checkbox"/> # Age	Continuous	Input	Read	

Default mode

☒ Read metadata ☐ Pass (do not scan)

Cancel Save

Auto Classifier Model

Projects / Classification of Diabetes / Classification Of Diabetes

auto

Run selection

Field Operations

- Auto Data Prep: Prepare your data for model building quickly and easily, without needing prior knowledge of the statistical concepts involved.
- Auto Classifier: Estimates and compares models to try out a variety of approaches for nominal and binary data.
- Auto Numeric: Estimates and compares models to try out a variety of approaches for a continuous numeric range.
- Auto Cluster: Estimates and compares clustering models that identify groups of records with similar characteristics.

Data Asset

Type

9 Fields

Partition

Diabetes

Diabetes

Outputs

Models

Outputs are available during your SPSS Modeler session. After your session ends, the outputs are no longer available.

☐ All results

☐ 9 Fields 6 minutes ago

Tabular View of Output

Projects / Classification of Diabetes / Classification Of Diabetes

View Model: Diabetes

Auto Classifier ②

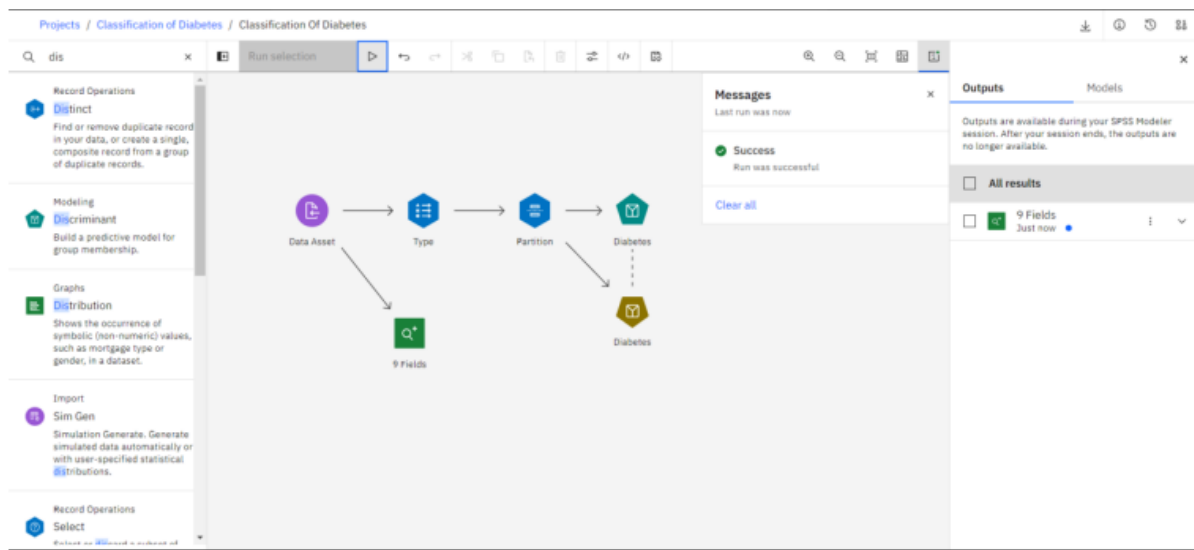
Models

Auto Classifier - Models ②

TARGET : DIABETES

USE	MODEL NAME	ESTIMATOR ▼	BUILD TIME (MINS)	NO. FIELDS USED	ACCURACY	ACCUMULATED ACCURACY	AREA UNDER CURVE	ACCUMULATED AUC	RECALL	PRECISION
<input checked="" type="checkbox"/>	Logistic regression.1	Nominal Regression	< 1	8	68.293	68.293	0.779	0.779	0.412	0.700
<input checked="" type="checkbox"/>	Discriminant.1	Discriminant	< 1	8	73.171	73.171	0.776	0.776	0.588	0.714
<input checked="" type="checkbox"/>	Tree-AS.1	CHAID	< 1	4	73.171	73.171	0.766	0.766	0.471	0.800
<input checked="" type="checkbox"/>	CHAID.1	CHAID	< 1	4	73.171	73.171	0.749	0.749	0.529	0.750
<input checked="" type="checkbox"/>	C5.1	C5.0	< 1	7	69.512	69.512	0.712	0.712	0.529	0.667

Discriminant Model



Model View

