

Exploratory Analysis

Catalogo de formulas quimicas para farmaceuticas e investigacion

Carbosynth

Analisis exploratorio de catalogo de formulas quimicas para su venta para diferentes sectores. Estas formulas quimicas contienen diferentes caracteristicas y propósitos tanto su precio a la venta.

Variable	Categorica	Numerica	Texto	Boolean	Presencia de nulos	Ejemplo
Substance Category	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	
Substance Fullname	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	
Experimental props	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	True , False
Bulk	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	True , False
Screening	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	True , False
Hazard property	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	True , False
Quantity	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	2,5,10,1000
Price	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	In USD
Quantity Category	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	Gram or Kg
Chemical Formula	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	
Molecular Weight	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	

Observaciones:

Algunos de los registros no tienen informacion de su formula quimica , ni de su peso molecular, por lo tanto no se obtiene informacion acerca de su propiedad de peligrosidad (Hazard Property).

Registros de formulas quimicas: 5663

Pasos realizados en esta etapa de analisis?

- Identificar tipo de variables
- Revision de presencia de nulos
- Encontrar un API o libreria para el parseo de XML
- Obtener por medio de REQUEST la formula quimica y peso molecular