#	Data produced by Portia's Instance-based Wrangler	Result after Data Cleaning using Trifacta-Wrangler
1	{"ranking":["#1"],"reviewCount":["1,951 Reviews"],"url":"https://www.tripadvisor.com/Attraction_Review-g304553-d500203-Reviews-Sri_Chamarajendra_Zoological_Gardens_Mysore_Zoo-Mysuru_Mysore_Karnataka.html","_type":"item","_cached_page_id":"92e6f61d91ab2a8084abd8c16746718f4dfc1c90","contact":["0821-2440752"],"things_to_do_in":["Mysuru (Mysore)"],"address":["Indiranagar | Near Mysore Palace , Mysuru (Mysore) 570 010 , India"],"_template":"f1b17f12b271f090607ff7b6e95e4c1a06353518","place":["Sri Chamarajendra Zoological Gardens,Mysore Zoo"]}	{"country":"India","attraction_in":"Mysuru (Mysore)","rank":"1","no_of_reviews":"1951","url":"https://www.tripadvisor.com/Attraction_Review-g304553-d500203-Reviews-Sri_Chamarajendra_Zoological_Gardens_Mysore_Zoo-Mysuru_Mysore_Karnataka.html","contact_no":"0821-2440752","address_info":"Indiranagar | Near Mysore Palace , Mysuru (Mysore) 570 010 , India","attraction":"Sri Chamarajendra Zoological Gardens,Mysore Zoo","no_of_attractions_per_location":29,"no_of_attractions_in_country":3251}
2	{"ranking":["#25"],"reviewCount":["679 Reviews"],"url":"https://www.tripadvisor.com/Attraction_Review-g187855-d247196-Reviews-Museo_Nazionale_del_Risorgimento-Turin_Province_of_Turin_Piedmont.html","_type":"item","_cached_page_id":"596264faf8959007afe032640a2aef8e0e8d0b57","contact":["+39 011 56 21 147"],"things_to_do_in":["Turin"],"address":["Via Dell'Accademia delle Scienze 5 | Piazza Carlo Alberto 8 , 10123 Turin , Italy"],"_template":"f1b17f12b271f090607ff7b6e95e4c1a06353518","place":["Museo Nazionale del Risorgimento"]}	{"country":"Italy","attraction_in":"Turin","rank":"25","no_of_reviews":"679","url":"https://www.tripadvisor.com/Attraction_Review-g187855-d247196-Reviews-Museo_Nazionale_del_Risorgimento-Turin_Province_of_Turin_Piedmont.html","contact_no":"+39 011 56 21 147","address_info":"Via Dell'Accademia delle Scienze 5 | Piazza Carlo Alberto 8 , 10123 Turin , Italy","attraction":"Museo Nazionale del Risorgimento","no_of_attractions_per_location":87,"no_of_attractions_in_country":10806}

Field Identified to be Transformed	Logic, Script used by Trifacta-Wrangler	Transformed Field
"ranking":["#1"]	**# token was removed to transform data-type of field from string to integer and records with empty rank were deleted** extract col: ranking on: `{digit}+` rename col: ranking1 to: 'rank' drop col: ranking delete row: empty([rank])	"rank":"1"
"reviewCount":["1,951 Reviews"]	**Extract number of reviews, removing ',' to fetch integer no of reviews, s set to 0 if value is empty** extract col: reviewCount on: `{number}` after: `"` before: `` rename col: reviewCount1 to: 'no_of_reviews' drop col: reviewCount replace col: no_of_reviews with: '' on: `,` set col: no_of_reviews value: 0 row: empty([no_of_reviews])	"no_of_reviews":"1951"
"address":["Indiranagar | Near Mysore Palace , Mysuru (Mysore) 570 010 , India"]	**Extract country from address, if unavailable mark as N/A** unnest col: address keys: '[0]' rename col: column_0 to: 'address_info' drop col: address extract col: address_info on: `{alpha}+` after: `{delim-ws}` before: `{end}` rename col: address_info1 to: 'country' set col: country value: 'N∨A' row: empty([country])	"address_info":"Indiranagar | Near Mysore Palace , Mysuru (Mysore) 570 010 , India" "country":"India"

	Aggregations across various dimensions performed, resulting in couple of aggregation tables. **1) To capture no of attractions in a location(using** attraction_in **field)** aggregate value: count() group: attraction_in rename col: row_count to: 'no_of_attractions' **2) To capture no of attractions in a country (using** country **field)** aggregate value: count() group: country rename col: row_count to: 'no_of_attractions_in_country'	"no_of_attractions_per_location": 29, "no_of_attractions_in_country":32 51
"contact":["0821-2440752"], "things_to_do_in":["Mysuru (Mysore)"] "place":["Sri Chamarajendra Zoological Gardens,Mysore Zoo"]	**Change field-names and Delete records where things_to_do_in is null** unnest col: contact keys: '[0]' rename col: column_0 to: 'contact_no' drop col: contact delete row: matches([things_to_do_in], `null`) extract col: things_to_do_in after: `"` before: `"` rename col: things_to_do_in1 to: 'attraction_in' drop col: things_to_do_in unnest col: place keys: '[0]' rename col: column_0 to: 'attraction' drop col: place	"contact_no":"0821-2440752" "attraction_in":"Mysuru (Mysore)" "attraction":"Sri Chamarajendra Zoological Gardens,Mysore Zoo"
"_type":"item" "_cached_page_id":"92e6f61d91ab2a 8084abd8c16746718f4dfc1c90" "_template":"f1b17f12b271f090607ff7 b6e95e4c1a06353518"	**Drop irrelevant auto-generated fields of Portia.** drop col: type drop col: cached_page_id drop col: template	
	Aggregation tables generated were Joined with other cleaned fields join with: no_of_attractions_per_location col: attraction_in = current_dataset.attraction_in, attraction_in1 = no_of_attractions_per_location.attraction_in, rank = current_dataset.rank, no_of_reviews =	"country":"India", "attraction_in":"Mysuru (Mysore)",

	current_dataset.no_of_reviews, url = current_dataset.url, contact_no = current_dataset.contact_no, address_info = current_dataset.address_info, country = current_dataset.country, attraction = current_dataset.attraction, no_of_attractions = no_of_attractions_per_location.no_of_attractions row: current_dataset.attraction_in == no_of_attractions_per_location.attraction_in

drop col: attraction_in1

rename col: no_of_attractions to: 'no_of_attractions_per_location'

join with: no_of_attractions_per_country col: country = current_dataset.country, country1 = no_of_attractions_per_country.country, attraction_in = current_dataset.attraction_in, rank = current_dataset.rank, no_of_reviews = current_dataset.no_of_reviews, url = current_dataset.url, contact_no = current_dataset.contact_no, address_info = current_dataset.address_info, attraction = current_dataset.attraction, no_of_attractions_per_location = current_dataset.no_of_attractions_per_location, no_of_attractions_in_country = no_of_attractions_per_country.no_of_attractions_in_country row: current_dataset.country == no_of_attractions_per_country.country

drop col: country1 | "rank":"1",

"no_of_reviews":"1951",

"url":"https://www.tripadvisor.com/Attraction_Review-g304553-d500203-Reviews-Sri_Chamarajendra_Zoological_Gardens_Mysore_Zoo-Mysuru_Mysore_Karnataka.html",

"contact_no":"0821-2440752",

"address_info":"Indiranagar | Near Mysore Palace , Mysuru (Mysore) 570 010 , India",

"attraction":"Sri Chamarajendra Zoological Gardens,Mysore Zoo",

"no_of_attractions_per_location": 29,

"no_of_attractions_in_country":3251 |