**Dataset Description (1.1)**

**Data Sources**

For this Rails project, I will be using the **Rick and Morty API** (https://rickandmortyapi.com/) to gather data about the Rick and Morty universe. This API provides three main endpoints that will serve as my data sources:

**1. Characters Endpoint (/api/character)**

* **Structure**: JSON objects containing character information
* **Key Attributes**:
  + Basic info: name, status (alive/dead/unknown), species, gender
  + Visual: image URL for character portraits
  + Location data: origin location, current location
  + Episode appearances: array of episode URLs
* **Volume**: 800+ characters available

**2. Locations Endpoint (/api/location)**

* **Structure**: JSON objects containing location/dimension information
* **Key Attributes**:
  + Basic info: name, type (planet, dimension, etc.), dimension
  + Residents: array of character URLs who live there
  + Creation metadata: when location was created
* **Volume**: 100+ locations available

**3. Episodes Endpoint (/api/episode)**

* **Structure**: JSON objects containing episode information
* **Key Attributes**:
  + Basic info: name, air\_date, episode code (S01E01 format)
  + Characters: array of character URLs who appear in episode
  + Creation metadata: when episode data was added
* **Volume**: 50+ episodes available

**Database Schema Design**

**Required Tables and Columns**

**Characters Table:**

* id (primary key)
* api\_id (original API ID for reference)
* name (string)
* status (string: "Alive", "Dead", "Unknown")
* species (string)
* gender (string)
* image\_url (string)
* origin\_location\_id (foreign key → locations)
* current\_location\_id (foreign key → locations)
* created\_at, updated\_at (timestamps)

**Locations Table:**

* id (primary key)
* api\_id (original API ID for reference)
* name (string)
* type (string: "Planet", "Dimension", etc.)
* dimension (string)
* created\_at, updated\_at (timestamps)

**Episodes Table:**

* id (primary key)
* api\_id (original API ID for reference)
* name (string)
* air\_date (string)
* episode\_code (string: "S01E01" format)
* created\_at, updated\_at (timestamps)

**Character\_Episodes Join Table (many-to-many):**

* id (primary key)
* character\_id (foreign key → characters)
* episode\_id (foreign key → episodes)
* created\_at, updated\_at (timestamps)

**Data Integration Strategy**

The three data sources are naturally interconnected through the API's reference system:

1. **Characters ↔ Locations**: Each character has an origin location and current location, creating two one-to-many relationships from locations to characters.
2. **Characters ↔ Episodes**: Characters appear in multiple episodes, and episodes feature multiple characters, creating a many-to-many relationship requiring a join table.
3. **Locations ↔ Episodes**: Indirectly related through characters, we can find which episodes feature characters from specific locations.

**Data Import Process**

The seed script will:

1. Fetch all characters from the API and populate the characters table
2. Extract unique location references and fetch location details
3. Fetch all episodes and populate the episodes table
4. Create character-episode associations based on API references
5. Establish character-location relationships for origin and current locations

This approach will ensure referential integrity while maintaining the rich relational structure needed for advanced Rails features like hierarchical navigation and complex searching.

**Database ERD (1.2)**

