

## Big Picture

### General Workflow

- Scrape Data
- Store data in SQL
- Perform Analysis on data
- Machine Learning Models
- Present data using flask and bootstrap website

### Who is this helpful for?

- Apartment Agencies
- People moving cities

### How do I provide Value?

- Better predictions
  - Better ML models
  - description/Amenities analysis
- gathering more comprehensive data then competition
  - City information
  - Information on listing agencies

## ERD

### Cities

- City ID (Primary Key)
- Name
- Cost of Living Indices (Groceries, Rent, Entertainment, Food, Healthcare)
- Population
- Age(f,l)
- Median rent
- Poverty rate
- Population density
- Most common industries\*
- Hospitals \*
- Airports \*
- schools\*
- Crime Rates \*
- Tax Rates
- Unemployment Rate
- Median Income
- Economic Growth Rate

### Apartment Complexes

- Complex ID (Primary Key)
- Name
- City ID (Foreign Key)
- Price Range
- Latitude
- Longitude
- Address
- Neighborhood
- User Reviews Average
- Amount of Units
- Building Age
- Type (e.g., High-rise, Gated Community)
- Safety Features (e.g., Security System Availability)
- Community Amenities (e.g., Pool, Gym)

### Apartment Units

- Unit ID (Primary Key)
- Complex ID (Foreign Key)
- Price
- Number of Beds
- Number of Baths
- Unit Amenities
- Square Footage
- Floor Level
- Renovation Status
- Lease Terms Flexibility

### Listing Agencies

- Agency ID (Primary Key)
- Name
- Reputation Rating
- Specialization

### Relationships:

- Cities to Apartment Complexes: One-to-Many (A city can have multiple apartment complexes, but each complex is located in one city).
- Apartment Complexes to Apartment Units: One-to-Many (An apartment complex contains multiple units, but each unit belongs to one complex).
- Apartment Complexes to Listing Agencies: Many-to-Many (Apartment complexes can be listed by multiple agencies, and each agency can list multiple complexes).

This relationship typically requires an associative entity to handle the many-to-many relationship.

- Apartment Complexes to Transportation Infrastructure: Many-to-Many (Complexes can be near multiple types of transportation infrastructure, and each infrastructure type can be near multiple complexes).

## Conversion

### Strong

#### Cities

- City ID (Primary Key)
- Name
- Cost of Living Indices (Groceries, Rent, Entertainment, Food, Healthcare)
- Population
- Age(f,l)
- Median rent
- Poverty rate
- Population density
- Tax Rates
- Unemployment Rate
- Median Income
- Economic Growth Rate

#### Leasing Company

- Leasing ID
- Name
- Contact

#### Apartment Complex

- Complex ID
- City ID →
- Name
- Url
- Price Range
- Lat
- Long
- Neighborhood
- User Reviews
- Unit Amount

- Type
- Details
- 

#### Apartment Units

- Unit Id
- Complex Id →
- MaxRent
- Num Beds
- Num Baths
- Square Footage

Weak

Multivalue

#### Complex\_Ammenities

- Complex ID→
- Amenity

#### Unit\_Ammenities

- Unit ID→
- Amenity

#### City\_schools

- City\_id →
- School

#### City\_hospitals

- City\_id →
- hospitals

#### City\_airports

- City\_id →
- airports

#### City\_crime

- City\_id →
- Assault
- theft
- Etc

#### City\_industries

- City\_id →
- Industry

Results of apartments

```
dict_keys(['listingId', 'resultPosition', 'favoriteState', 'hasRentSpecials', 'phoneNumber',
'listingTypeId', 'listingStatusId', 'deferCarouselLoading', 'carouselCollection', 'galleryCollection',
'mediaCollection', 'listingCity', 'listingState', 'listingZip', 'listingAddress', 'listingCountry',
'listingDMA', 'listingName', 'listingMinRent', 'listingMaxRent', 'AnalyticsAdLevel', 'location',
'rentals', 'reviews', 'isOpenLeadForm', 'searchCriteria', 'geo', 'defaultZoomLevel', 'isTierTwo',
'profileType', 'languagesSpoken', 'isMF', 'isFromFeed', 'feedType', 'propertyType',
'listingNeighborhood', 'listingCounty', 'companyKey', 'minBeds', 'myApplicationsUrl', 'conduit',
'reviewActions', 'abData', 'tourMediaId', 'videoMediaId', 'has3DTour', 'hasVideo', 'imageCount',
'videoCount', 'photoCount', 'virtualTourCount', 'hasTCPA', 'hasIntegratedTours',
'hasPropertyWebsiteLink', 'engrainKey', 'todaysOfficeHours', 'listingSpecialties',
'listHubListingUri', 'ip', 'leasingOffice', 'nextOpenOfficeHours'])
```

## Results of Zillow

```
dict_keys(['regionIds', 'breadcrumbs', 'buildingName', 'streetAddress', 'latitude', 'longitude',
'aerialViewVideoId', 'streetViewTileImageUrlLocationModuleLatLong',
'streetViewTileImageUrlLocationModuleAddress', 'lotId',
'streetViewMetadataUrlMediaWallLatLong', 'streetViewMetadataUrlMediaWallAddress',
'mapTileGoogleMapUrlLocationModule', 'mapTileGoogleMapUrlFullWidthMax', 'address',
'streetViewTileImageUrlHalfWidthLatLong', 'streetViewTileImageUrlHalfWidthAddress', 'city',
'state', 'zipcode', 'isWaitlisted', 'contactInfo', 'isLandlordLiaisonProgram', 'buildingType',
'instantTourTypes', 'isInstantTourEnabled', 'bestGuessTimezone', 'rentalInstantTour',
'photoCount', 'amenitiesVRModels', 'videos', 'galleryPhotos', 'galleryAmenityPhotos',
'fullAddress', 'bdpUrl', 'zpid', 'buildingPhoneNumber', 'county', 'listingFeatureType',
'ungroupedUnits', 'floorPlans', 'metaData', 'buildingAttributes', 'nearbyCities',
'nearbyNeighborhoods', 'country', 'nearbyZipcodes', 'nearbyBuildingLinks', 'amenitySummary',
'comps', 'marketInsightsUrl', 'isLowIncome', 'isSeniorHousing', 'isStudentHousing', 'ppcLink',
'neighborhood', 'reviewsInfo', 'housingConnector', 'screeningCriteria',
'buildingRentalPremiumPackagesInfo', 'assignedSchools', 'nearbyAmenities', 'walkScore',
'transitScore', 'bikeScore', 'description', 'amenityDetails', 'communitySubType', 'homeInsights',
'bestMatchedUnit', 'photos', 'amenityPhotos', 'staticMapSatellite', 'streetViewLatLong',
'streetViewAddress', 'staticMap', 'thirdPartyVirtualTours', 'currency', 'detailedPetPolicy',
'specialOffers', 'listingMetadata', 'commonUnitAmenities', 'vaLoanStatus', 'engrain', 'homeTypes',
'providerInfoList', 'unitOverview', 'unitsVRModels', '__typename'])
```