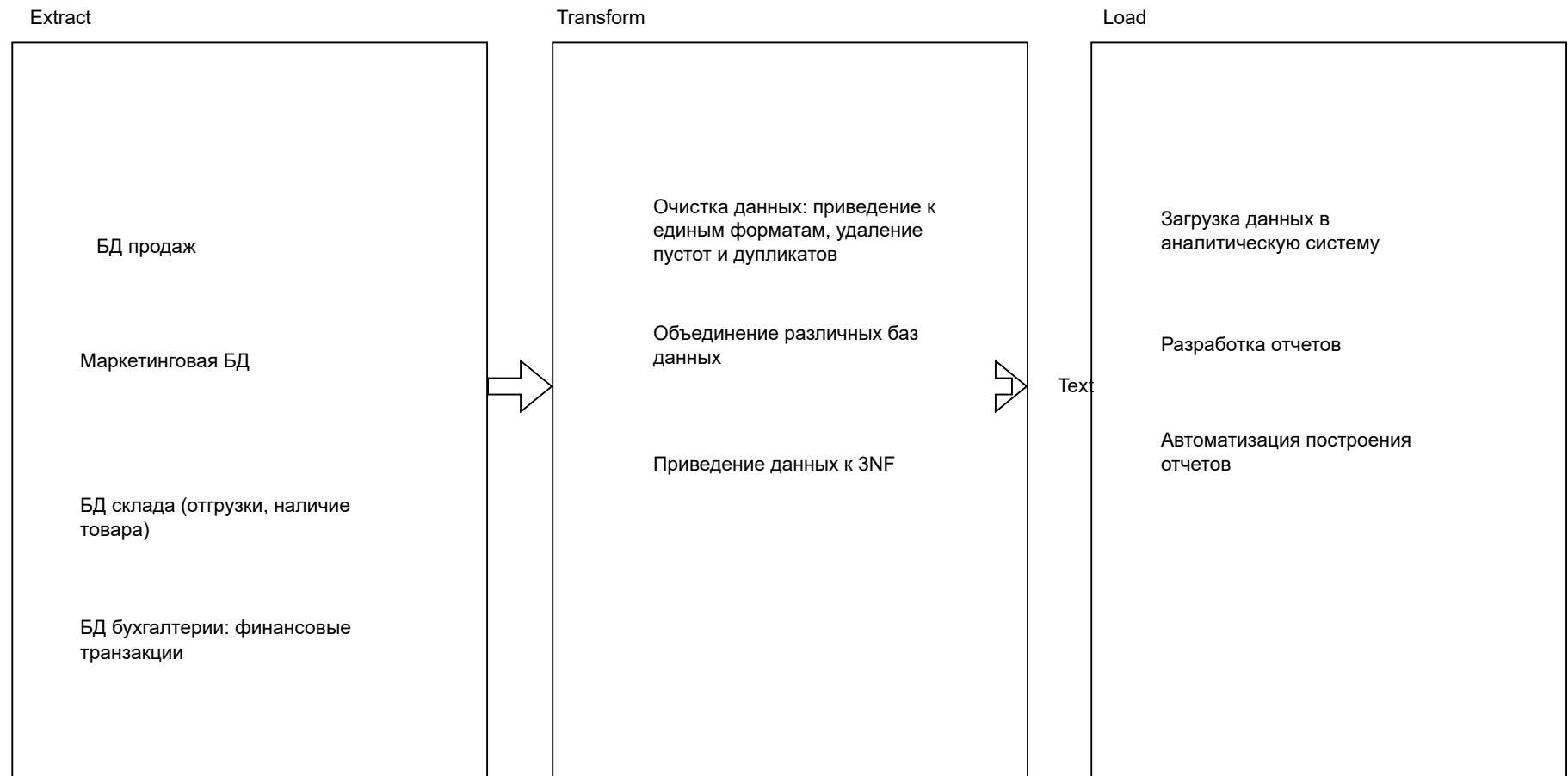


Архитектура ETL процесса для сбора и анализа данных компанией которая хочет провести маркетинговую кампанию



Архитектура достаточно типична и описывает общий подход к подготовке данных для анализа маркетинговой компании. Потребуется данные из разных источников - следовательно, потребуется их объединение (кажется, это называется "проброс айдишника", когда по одному ключу можно искать данные пользователя в разных базах данных).

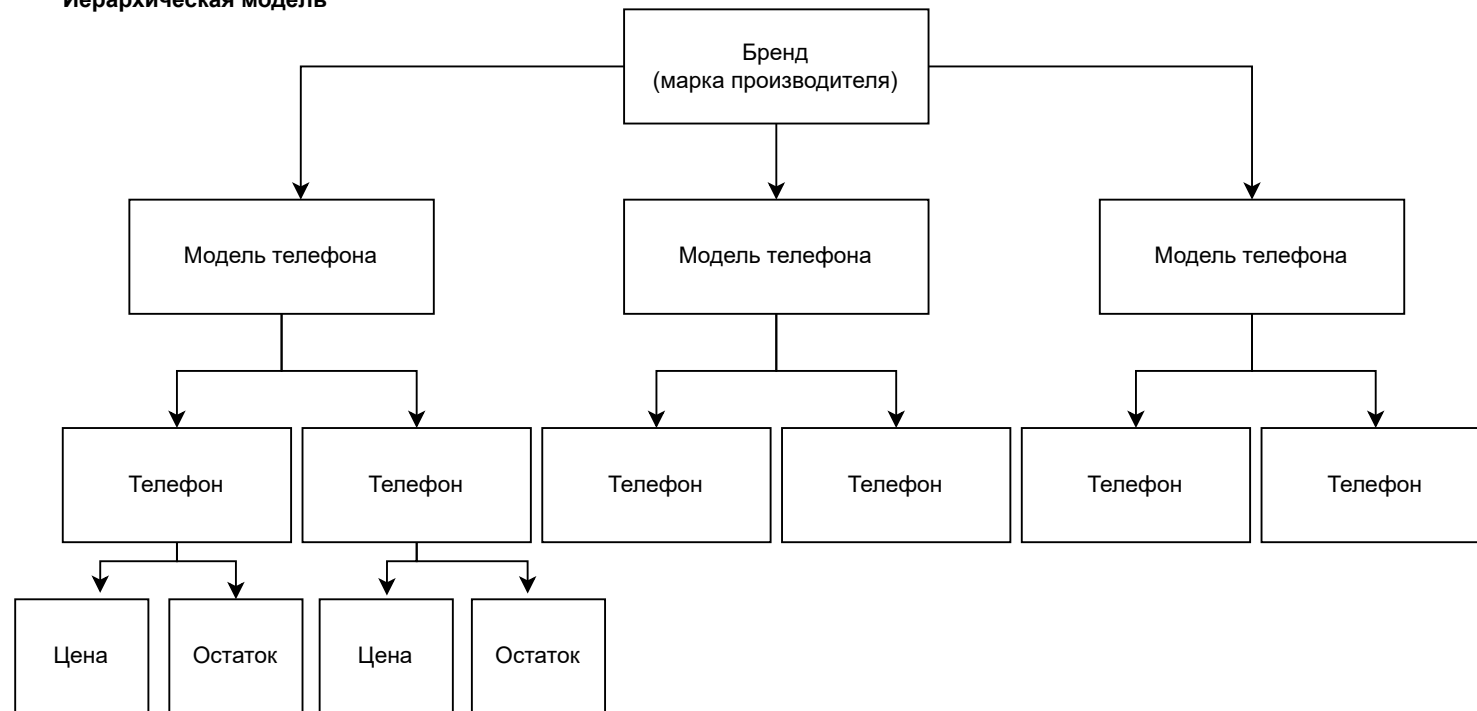
В любом случае потребуется очистка данных от дубликатов, пустых строк, ошибочных данных. Потребуется приведение данных к 3NF - чтобы можно было работать дальше.

Результат усилий потребуется сохранить в аналитической БД и подключить к ней систему визуализации (например, Power BI).

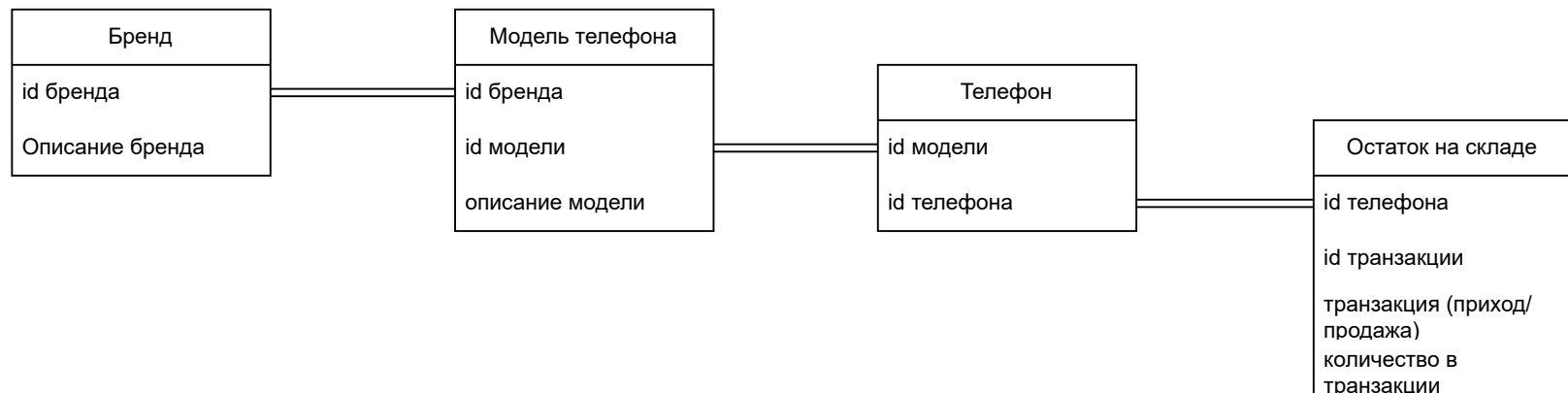
В итоге мы хотим получать актуальные переработанные данные в виде отчета о маркетинговой компании.

Построить реляционную и иерархическую модели данных для магазина который продает телефоны

Иерархическая модель



Реляционная модель



Исходная таблица. Все строки атомарны и уникальны: имеем 1NF

Employee_ID	Name	Job_Code	Job	City_code	Home_city
E001	Alice	J01	Chef	26	Moscow
E001	Alice	J02	Waiter	26	Moscow
E002	Bob	J02	Waiter	56	Perm
E002	Bob	J03	Bartender	56	Perm
E003	Alice	J01	Chef	56	Perm

2NF

Employee_ID	Name	City_code	Home_city
E001	Alice	26	Moscow
E001	Alice	26	Moscow
E002	Bob	56	Perm
E002	Bob	56	Perm
E003	Alice	56	Perm

Employee_ID	Job_Code	Job
E001	J01	Chef
E001	J02	Waiter
E002	J02	Waiter
E002	J03	Bartender
E003	J01	Chef

3NF

City_code	Home_city
26	Moscow
56	Perm

Job_Code	Job
J01	Chef
J02	Waiter
J03	Bartender

Employee_ID	Name	City_code
E001	Alice	26
E002	Bob	56
E003	Alice	56

Employee_ID	Job_Code
E001	J01
E001	J02
E002	J02
E002	J03
E003	J01