| **№** | **Теория** | **Теоретические задания** | **Практические задания** |
| --- | --- | --- | --- |
| **№** | **Теория** | **Теоретические задания** | **Практические задания** |
| 1 | Нормализация реляционной БД. Основные нормальные формы: 1НФ, 2НФ, 3НФ, НФБК. Дополнительно изучить: 4НФ, 5НФ. 6НФ. Якорная модель. | Что такое транзитивная связь | Привести таблицу Sell\_Out к 3-й нормальной форме |
| 2 | Типы связей, один к одному, многие ко многим т.д. |  |  |
| 3 | Денормализация реляционной БД. |  |  |
| 4 | Понятие "Хранилище данных" | Какие задачи решает КХД, DWH |  |
| 5 | OLAP, OLTP, в чём отличие и почему OLTP не подходит для аналитических систем | Преимущества OLAP, преимущества OLTP. Примеры где используются такие модели |  |
| 6 | Что такое ETL и для чего он нужен. Основные проблемы. Отличие ETL от ELT подхода. | Преимущества ETL, преимущества ELT |  |
| 7 | Dimension model. Какие данные должны хранится в Dimension. | Что такое дименшн, как определить. Подумать, что из столбцов в Sell\_Out можно выделить в дименшн. |  |
| 8 | Fact, какие данные должны хранится в Fact. | Что такое факт, как определить. Подумать, что из столбцов в Sell\_Out можно выделить в факт. |  |
| 9 | Меры (measures), типы мер. | Агрегируемые (аддитивные), не агрегируемые, полу-агрегируемые . Привести примеры | Определить меры в таблице Sell\_Out |
| 10 | Гранулярность данных. |  | Добавить в таблицу Sell\_Out столбец CCD - Consistent Change Data. В нём должна быть конкатенация всех столбцов через разделитель "#~#", кроме мер (measures). Понять гранулярность существующих данных в таблице. |
| 11 | Что такое суррогатные ключи, для чего нужны и какие приемущества они приносят в DWH. | Откуда/как берутся сур ключи. Как они связаны с натуральным (бизнес) ключом (тип связи). |  |
| 12 | Принцип построения звезды и снежинки, различия и недостатки. |  | На бумаге выделить из таблицы Sell\_Out дименшены и факт. Разнести столбцы по таблицам, показать связи - нарисовать звезду. |
| 13 | SCD, зачем нужно и какие виды бывают. |  |  |
| 14 | Инкрементальный подход загрузки данных в DWH. | Что это, зачем нужно. Способы реализации, примеры. |  |
| 15 | Витрина данных. Зависимые и независимые витрины. | Что это, зачем нужно. |  |
| 16 | Какие области хранилища существуют и для чего каждая из них нужна | Staging Area (STG), Operational Data Storage (ODS) etc |  |