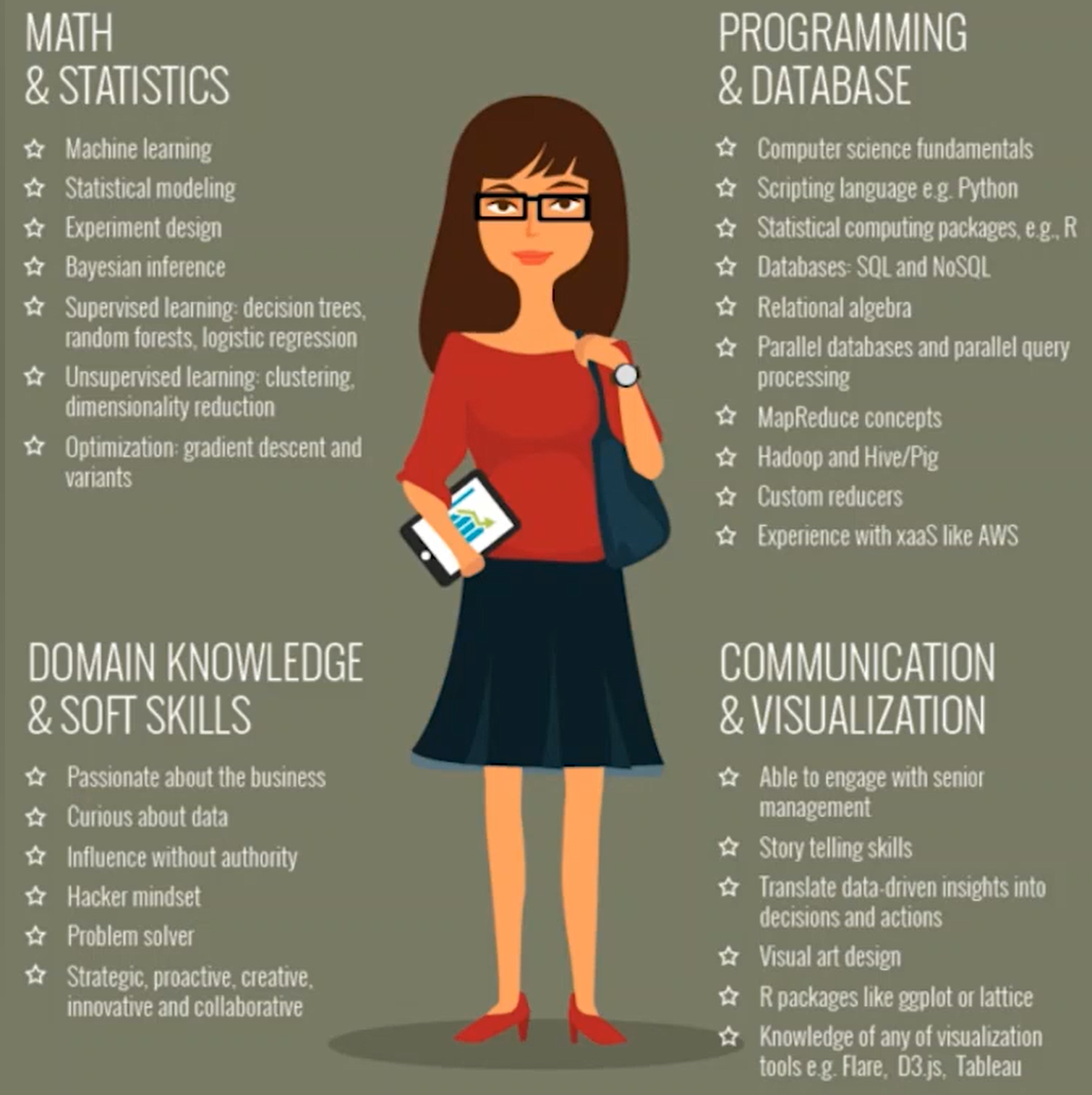
При написании данной дипломной работы я ставил перед собой следующие задачи: как устроены модели машинного обучения, как происходит процесс обучения моделей, зачем это нужно, какие ключевые концепции лежат в основании процессов, связанных с машинным обучением и какие экономические задачи можно решать, используя машинное обучение.

Методы помогают нам обрабатывать и преобразовывать данные. Атрибуты помогают нам увидеть, что хранится в данных.

DataFrame в Pandas – это не просто таблицы, но и множество атрибутов и методов к ним, состоящие из множества Series (векторов-столбцов).



Бизнес-аналитик формируем проблемный вопрос, который нужно решить и собирает требования. Дата инженер собирает и обрабатывает данные. Дата Сайентист анализирует и подбирает модель. Разработчик машинного обучения внедряет модель в программу.

Для начала данные нужно подготовить (чистка пустых данных), формируем признаки (фичи) датафрейма, агрегируем и т.д. Далее с помощью методов машинного обучения обучаем нашу модель и прогнозируем целевую переменную. Затем мы валидируем нашу модель, проверяем, что она предсказывает верно при работе с тестовыми, а не тренировочными данными.

Подбор коэффициентов b0 и b1 – это задача машинного обучения.

Задача всего обучения – научить модель находить общую ЗАКОНОМЕРНОСТЬ в данных, обучивших на тренировочной выборке, в которой есть некоторая частная закономерность.