データ・トレーニングスキル一覧

Googleスプレッドシートに関するスキル**:**

* データのインポート、ファイル形式の違い、ExcelとGoogleスプレッドシートの使い分け
* 基礎的な能力:
  + 検索
  + ソート
  + フィルタリング
  + ペインの固定
  + 平均値、中央値、合計、最大値、最小値や変化率といった基本的な数学の知識
    - $で数式（中のセル）を固定する
  + セルと列のフォーマット
    - 条件付き書式
* ピボットテーブル
* データクリーニング
  + 列の追加と削除
  + 特殊なペースト
  + Transpose（行と列を入れ替える関数）
  + ワイルドカードを含む検索と置換
  + Trim（不要なスペースを削除する）やproper（先頭の文字を大文字にする）などの関数
  + 列を分割して情報を抽出する: left/right/mid
  + 日付と時刻の操作
* データを結合したり、クリーニングしたりする時に使うVlookup
* データを分類し、整理するために使うIf/then

**その他のスキル:**

* （仕事の）段階やあなた自身を検証する方法を記録するなど、再現可能な調査を行う
* データ分析について質問したり、助けてくれたりする場所（を持つ）
* Cometdocs やTabulaでデータをPDF化する
* スプレッドシートを使って（データを）整理された状態に保つ

**データセット:**

* 新しいデータセットについてまず調べるべきこと
* 国勢調査のように、全ての記者が知っておくべきデータセットとその使い方
* スクープに重要となるデータセット
* データ辞書とその使い方
* Fipsのような一般的なエンコーディング
* スクレイピングとは何か;プログラミングで何ができるのか
* あなたのデータセットを作成する際のベストプラクティス

**ストーリーにデータを役立てる:**

* データストーリーの七つの型
* データからストーリーを発見する
* あなたの仮説を数字で裏付ける — あなたのストーリーで数字に置き換えられるものは何か?
* オルタナティブ・ストーリーにデータを役立てるやり方
* グラフィック編集者との協働
* 倫理と [バッドデータへの対処](https://github.com/Quartz/bad-data-guide)
  + 相関すなわち因果関係ではない
  + チェリー・ピッキング
  + コンテクストの重要性
  + 誤差、調整と未調整