データ・チートシート: 基礎編

|  |  |
| --- | --- |
| **Googleスプレッドシートの長所** | **Excelの長所** |
| クラウドストレージ(Wifiを搭載した任意のデバイスからアクセスできます) | プライバシーとセキュリティ |
| 最大 200 万個のセルを保存できます | 大量のデータを分析可能 |
| コラボレーションと共有 |  |
| 自動保存とバージョン履歴 |  |
| 無料 |  |
| 他のプログラム(ドキュメント)や異なるコンピュータでも表示を維持します |  |

**データのレベル:**ワークブック (ファイル) -> シート (画面上で表示しているもの) -> 行 + 列 -> セル

**セルの書式設定:**セルまたは列（に入力した数値の）表示方法を制御します。たとえば、日付 (2018/02/26 または 02/26/18もしくは 2 月 26 日など) や、小数点以下の桁数などです。

* ツールバーの書式をクリックし、番号 (ドロップダウンの最初の選択肢)に進み、次のポップアウトで書式を選択します。

**[ペインの固定]:**スクロールしても画面から消えないよう、行または列 (ふつう最初の行)をロックできます。

* 大きなデータ・セットを扱う場合に便利です。
* 表示 ->固定 -> 1 行目の順に選択し、１行目をフリーズします。

**検索:**スプレッドシート内の特定の値を検索します。

* 他のアプリケーション同様Ctrl + f でもいいですし、ツールバーの編集 -> 検索と置換でもいいです。
* 探している値をハイライトで強調し、スプレッドシートに存在する値の数を検索バーに表示します。（検索バーの横の）矢印で、ハイライト表示された値間を移動できます。

**並べ替え: データセット**内の最大値または最小値を確認するか、日時またはアルファベット順に行を配置します。

* スプレッドシートの左上隅をクリックして、すべての行と列を選択します。
  + この手順をスキップすると、一つの列だけを並べ替えてしまうので、データがごちゃちゃになります。
* データ -> 範囲を並べ替え ...
* [データにヘッダー行が含まれている]のチェック ボックスをオンにします。
* 並べ替えの基準となる列を選択します。
* A-Z (数値の場合は昇順)または Z-A (降順)
* 並べ替えの基準となる別の列を追加するには、[次に] を選択します (多くの行が同じ値を持つ場合に便利です)。

**フィルタリング:**列の値に基づいて、データをフィルター処理します。

* フィルタにはオンかオフの状態があります。ツールバーまたはデータ -> じょうごの形をした「フィルタを作成」で取得します。
* フィルタがオンならば、ヘッダー行の横に（三角の形をした）線が表示されます。これをクリックすると、その列の値を用いてフィルタリングされます。
* アクティブなフィルタがある場合は、列のタイトルの隣にフィルタアイコンが出ます。
  + **条件付きフィルタリング:**同じ手順を踏みますが、フィルタする列をクリックすると条件付きフィルタリングを選択できます。「10 以下の値を持つ行のみ」など値の範囲によってフィルタリングできます。
* （設定済みの）フィルタを頻繁に使う時は、データ -> フィルタ表示-> フィルタ表示として保存する　を選択してください。

**数式:**セルの値を計算します。たとえば、セルに =2+2 と入力すると、セルには 4が表示されます。

* 数式は何百種類もあります。これらはすべて = で始まり、通常は () に対象のセルまたは範囲が含まれます。
* 例: =AVERAGE (C1:C22) は、C列の１行目から 22 行目までの値の平均を計算します。
* よく使われる数式:
  + =MEDIAN()：データセットに外れ値が含まれているとき、平均の代わりに使います
  + =MIN()：データセット内の最小値
  + =MAX() ：データセット内の最大値
* 変化率: ((value2-value1)/value1)\*100
  + Googleスプレッドシートでは、=((A2-A1)/A1)
  + それから、セルの表示形式をパーセントに変更します。
* $ を使用すると、別のセルに数式を適用した場合でも、値が固定されます。
  + =$A1 は、数式を左右（のセル）に適用しても A (列) を参照します。
  + =A$1は、数式を上下に適用しても1 (行目) を参照します。
  + =$A$1 は（行と列の）両方を固定するので、シートを上下左右に移動しても、常にA1 を参照します。

**平均/代表値と中央値:**平均はすべてのデータを足し合わせ、値の個数で除算します。すべての値が共線であるといえるとき、中央値は分布の中央に位置します。外れ値が存在する場合は、一般に中央値を使用します。それ以外の場合は、平均を使用します。