Why Computing for Engineers?
Problem solving

Basic Spreadsheet Concepts
- Why Use a Spreadsheet?
- Key Concepts
- Charts
- Organising Data: sorting and filtering
- Useful functions and Features
Computers have changed engineering ...

Computers have changed engineering ...
How computing is used in engineering?

• **Automation** is a major application of computing in engineering
  – There are many other applications of computing in engineering. More to come.
  – Message: *Computing will play a key role in addressing grand challenges in engineering, e.g. aging infrastructure, etc.*
  – [http://www.engineeringchallenges.org](http://www.engineeringchallenges.org)

• Automation: Computers/machines repeatedly performing the same procedure
  – Procedure: a sequence of *instructions*

• Computing skill: The ability to give instructions to computers to perform the intended tasks
How computing is used in engineering?

- *Analysis* and management of often huge datasets, e.g. production line, climate trends, mining, etc.

- *Modelling* of real or planned systems, e.g. structural analysis, complex interconnected processes, etc.

- *Tools* to support good engineering practices, e.g. project and resource management, etc.
Problem Solving

• Preparing to Solve a Problem
  – What is the overall purpose of the problem (define)?
  – What information is known (input)?
  – What information must be determined (output)?
  – What fundamental principles apply to the problem?
  – What will be the overall solution strategy?
  – How can we check correctness our solution?
  – How can we test our solution(s)?
Problem solving

Engineering:

invention + problem solving + ...

• How can you use computers to solve problems for you?
  – How can you use computers to understand, investigate, test and design?

• We will use a motivating example (next page) and learn how spreadsheet can help us to do some problem solving
Counting heart beat automatically

- Pulse oximetry sensor

- We will use spreadsheet to help us to understand how to count heart beat automatically

http://pulsesensor.com
Why Use a Spreadsheet?

Many things that can be done with a pencil, a pad of paper, and a calculator can be done much faster, far more accurately and conveniently using a spreadsheet.

The biggest advantage of using spreadsheet is that every time a user makes changes, the spreadsheet automatically recalculates all the other related values.

Nowadays, spreadsheets also allow easy creation of charts and offer useful statistical and mathematical functions.

The best known commercial spreadsheet application is Microsoft’s Excel®, but there are also free, open-source applications with similar functionality and even partial compatibility, notably OpenOffice.org™’s Calc, LibreOffice Calc (which shares an ancestor with Calc), Gnumeric from the long-running GNU project and the online Google Sheets.

In this course we use OOo Calc because it demonstrates the features of a good spreadsheet application, and is free. Since Excel is more widely used in industry, differences between the two applications will be noted periodically.
Each sheet is divided into rows and columns.

Rows are labelled as integers starting from 1 (up to 1 million+); column labels are from A to Z, AA to AZ, BA to BZ, etc. (up to 1000+)

The intersection of a row and column forms a cell. Each cell has a unique cell reference. For example, the cell at the intersection of column D and row 10 is called D10.
Every cell in a spreadsheet contains either a **constant** or a **formula**.

A **constant** is an entry that the spreadsheet does not change, for example, test marks, student name or a date.

A **formula** is a combination of numeric constants, cell references, arithmetic operators, and functions that returns the result of a calculation. Formulas are prefixed by an equal sign (=).
Functions in Formulas

A function performs a predefined computational task

For example the function \texttt{AVERAGE(B2:B6)}, in cell B8, calculates the average of all the cells in the range from B2 to B6 inclusive.

Calc provides many useful functions for mathematical, statistical, financial and other tasks. Documentation can be found at

https://wiki.openoffice.org/wiki/Documentation/How_Tos/Calc:_Functions_listed_by_category
Copy and Paste Commands

The **copy** command duplicates the contents of a cell, or range of cells.

The **paste** command copies the contents to the destination cell, or range of cells. However, a formula is *not copied exactly*, but is adjusted as it is copied, depending on the destination cell.

For example, if the formula in cell **B8** is copied to cell **C8**, it is adjusted so that the cells referenced in the new formula are in the same *relative position* as those in the original formula.

- Original formula is: \( =\text{AVERAGE}(B2:B6) \)
- Copied formula is: \( =\text{AVERAGE}(C2:C6) \)

Same principle applies to:
- inserting rows and columns
- deleting rows and columns
- filling operations
Filling combined copy and paste over a range of cells

Step 1: select range, usually in one direction, down or right

Step 2: pick the *Edit – Fill* menu item to fill *down* (if vertical) or *right* (if horizontal)

(Calc also supports fill left and fill up)

**Auto-fill** shortcut:

Step 1: select initial cell (the one with the original formula, B8)

Step 2: grab the auto-fill handle (square dot at bottom right)

Step 3: cursor changes to +, drag the range, fill is completed on release

Extra-special deal: select two adjacent cells containing numbers, auto-fill completes an arithmetic progression with tooltip feedback
Absolute vs Relative References

In a formula, if the cell reference has $ signs in front of the row and column designation, it is considered an absolute reference and the cell reference remains unchanged throughout all copy/paste operations, e.g., $B$10.

A relative reference, on the other hand, adjusts during copy operations and is specified without $ signs, e.g., B2.

A mixed reference uses a single $ sign to make the column or the row absolute, leaving the other as relative. For example F$6 or $F6.
A special form of absolute reference is to give a cell a **name**, using letters only.

Select a cell (or a **range**) and enter the intended name in the **name box**, or pick **Insert – Names – Define** from the menu (lists all existing names).

Names can be used in formulas (press Enter to accept suggestions)
Formatting Cells

The appearance of a cell can be changed by altering fonts, borders, colour fill and number formatting, including indents.

Note that changing the format of a number affects the way the number is displayed but *does not change its value*.

Formats are set using either presets on the toolbar:

- only has currency, percent, standard (General)
- allows change of precision for fixed decimal-place numbers
- does not allow change of precision for % or scientific notation

... or by using the Properties sidebar on the right of the sheet*.

* or menu item Format – Cells for earlier Calc versions.
Formatting Cells

Some of the value formats available in Calc are:

**General format** *(default)*

A number is displayed the way it was originally entered
For example, 247 or 247.58 or 2257.42E+10 (but 123.00 becomes 123)

**Number format** *(default right-aligned)*

125.345 or 10,200.45 or 10200.45 or -135.50 or 1.355e+2 etc
(unfortunately scientific format precision is set very clumsily with the Properties sidebar: more options icon, custom format, enter more zeroes)

**Percentage format**

23% or 23.56% or 9% (remember 9% is stored as 0.09, not 9)

**Date format**

May 2, 2015 or 2/5/2015 or 02-May-15, etc
(dates are stored as fractional days since 30-Dec-1899, allows dates to be subtracted; time = portion of a day so 0.75 represents 18:00:00 or 6pm)

**Text format** *(enter numeric strings by prefixing with a single quote, e.g., '001)*

text strings are left-justified, even of they look like numbers
Calc Concepts

- Key knowledge and skills:
  - Menus, spreadsheet workbook and sheet structure
  - The active cell
  - Formulas and addressing
  - Selecting cell ranges, copy/paste
  - Filling down and right using menu or auto-fill handle
  - Named cells
  - Display formats
Charts

The graphical representation of data can be an attractive, easy-to-understand way to convey information.

Any good spreadsheet application will let you create charts from the spreadsheet data, with just a few simple keystrokes or mouse clicks.

<table>
<thead>
<tr>
<th>A</th>
<th>B</th>
<th>C</th>
<th>D</th>
<th>E</th>
<th>F</th>
<th>G</th>
<th>I</th>
<th>J</th>
<th>K</th>
<th>L</th>
<th>M</th>
</tr>
</thead>
<tbody>
<tr>
<td>Student</td>
<td>Test1</td>
<td>Test2</td>
<td>Final</td>
<td>Average</td>
<td>Weighted Total</td>
<td>Grade</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Susie</td>
<td>100</td>
<td>85</td>
<td>81</td>
<td>88.67</td>
<td>87.50</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Jun Wai</td>
<td>75</td>
<td>87</td>
<td>92</td>
<td>84.67</td>
<td>85.90</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>John</td>
<td>85</td>
<td>92</td>
<td>77</td>
<td>84.67</td>
<td>82.40</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Masood</td>
<td>94</td>
<td>62</td>
<td>76</td>
<td>77.33</td>
<td>78.60</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sandeep</td>
<td>64</td>
<td>98</td>
<td>83</td>
<td>81.67</td>
<td>80.30</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

![Charts Image]

If the corresponding data values were changed, the application would automatically redraw the charts, based on the new values.
Creating Charts and Plots

You can produce different types of charts with Calc, including column and bar charts, and line and scatter plots (more applicable to engineering data).

Procedure: select the source data range, (usually including headers), click on the Chart Wizard icon.

Title and axis labels are at the Chart Elements step, but can add or edit later.
Chart Adjustments – Colours

- The default chart formats are only a *starting point*. Don’t rely on them completely, adjust the presentation in order to improve its effectiveness as a way of communication information, not just looking nice.

  Bright *fill colours* distract the eye away from the relationship.

  So do bold (saturated) colours, this is the Calc default 😒

  Use neutral/pastel colours and add a thin border (can set using Tools – Options - Chart).

  A dark palette is also OK, no border needed

*Chart area* should have an outline to show against background cells (see first chart example, default is white background and no border)

*Position* may be locked to cell (pick Format – Anchor, select *To Page*)
Chart Adjustments

- Adjustments are made in Edit mode (double-click chart), then double-click on item (data series, axes, chart area), or pick from the Format menu
  - Set *appearance* with Borders (outline styles), Area (fill colour) and Transparency
  - Set *axis limits* and precision on *Numbers* tab if auto is unsuitable (as it often is if the data format is not General)
  - Edit *title* and *axis labels*
  - Remove *unwanted data series* from the original source range using Format – Data Ranges (Data Series tab)

- You can find more information on all of these from *Calc Help*, and also from *online tutorials* on Calc (available on the class web page)

- Demo sheets in lecture spreadsheet
  - *ASX*: draw line chart of stock market trends
  - *Trajectory*: scatter plot of a ballistic trajectory with smooth lines called splines. A good example, as the formulas are linked to named cells that can be easily changed in one place.
Beware Chartjunk

- Default spreadsheet charts (especially Microsoft Excel) are overloaded with many fancy features that can obscure rather than reveal data relationships.
- This kind of useless embellishment is called **chartjunk**, a term coined by Edward Tufte (*Envisioning Information*, 1990).


**Excel column chart defaults**

What the user wanted

**Calc’s default line chart has fat point markers and no lines**
How to Lie with Charts

The prize for the most useless and misleading chart type goes to the **pie chart**, and especially if it has 3D effects.

*Distribution of 5 candidates’ votes at 3 polling places*

Organising Data – Sorting

• Data values can be sorted in ascending or descending order, using numeric or text comparisons as appropriate

• Procedure:
  – Select range **must include all associated values**: select complete rows (though headers won’t be recognised 😞)
  – Pick Data – Sort
  – Select columns, or Options if headers don’t appear
  – Can have multiple criteria, ascending or descending
  – Destination can be a different range
  – Formulas are adjusted

![Sort criteria options](image)
Organising Data – Filtering

Filtering is a way of displaying rows that contain values you want to work with, and temporarily hiding those that are not relevant at the moment.

There are three kinds of filter in Calc, in increasing complexity:

- **AutoFilter**, which includes individual selection and top 10
- **Standard Filter**, which allows combinations of conditions
- **Advanced Filter**, for more complex criteria (not covered here)

After selecting columns, or a range with headers, **Data – Filter – Autofilter** creates a filter drop-down option against each column.

Filters are successive, but Top 10 is global, so this shows which of the overall top 10 profit records are XYZ products from the East region, *not* the top 10 profits for this product and region.
Organising Data – Standard Filter

The **Standard Filter** dialogue box can be picked from any of the filter drop-downs or the **Data – Filter** menu (cancels any current autofilter selections)

- Applies conditions from any column, not just single values

![Standard Filter Dialogue Box](image)

- However, conditions are prioritised with AND binding more tightly than OR, so can’t filter (say) Quantity < 1000 AND Product is either ABC or XYZ (or any other alternatives)

- Filters are easily lost (remove autofilter, rows remain hidden but criteria disappear), requires **Format – Row – Show** to fix.

- Excel supports multiple values in its autofilter, much more useful
Predefined Functions

- Calc offers a wide range of mathematical and related predefined functions for use in formulas. Where appropriate they apply to ranges as well as single values. They can be entered in UPPER or lower case.

- Brief list of functions by category:
  - Elementary functions: sum, max, count, mod, round, abs
  - Engineering mathematics: pi, exp, log10, sqrt
  - Trigonometry: sin, cos, atan2; hyperbolic trig sinh, sech, acoth
  - Advanced mathematics: bessel functions, complex numbers
  - Statistics: many
  - Logical or conditional: if, and, or, not, countif, sumif
  - Unit conversion: convert_add, radians, degrees
  - Financial: fv, opmt, db for modelling investments, loans, depreciation etc
  - Text: len, trim, mid, substitute, concatenate, upper/lower
  - Spreadsheet: match, lookup, vlookup

- The Function wizard is helpful in composing the formula.

- Documentation:
  https://wiki.openoffice.org/wiki/Documentation/How_Tos/Calc:_Functions_listed_by_category
Arithmetic Functions

Common Functions:

- **SUM**
  - Example: SUM(C2:C8) equals 94

- **AVERAGE**
  - Example: AVERAGE(C2:C8) equals 13.428571

- **COUNT**: Counts the number of cells that contain numbers
  - Example: COUNT(B2:B8) equals 4

- **COUNTA**: Counts the number of cells that are not empty
  - Example: COUNTA(B2:B8) equals 6
Conditional Accumulation Functions

- **COUNTIF**: Counts the number of cells within a range that meet the given criteria
  - Example: **COUNTIF**(C2:C8; ";<10") equals 2

**Important note**
Function arguments in Calc formulas *(arguments are the values inside the parentheses)* are separated by a *semicolon* and not a *comma*. This differs from almost every other functional application in existence, including Excel, Matlab, and most mathematical and programming notations.
Using the Function Wizard

- **SUMIF**: Adds the cells specified by a given criteria.
  - Example: `SUMIF(B2:B8; "Absent"; C2:C8)` equals 29

If omitted, the third argument to `sumif` defaults to the first range, but here it’s column C not B
Conditional Functions

- **IF**: Returns one value if a condition you specify evaluates to TRUE and another value if it evaluates to FALSE.

\[
\text{IF}(\text{logical_test} ; \text{value_if_true} ; \text{value_if_false})
\]

Examples:

\[
\text{IF}(C5<50 ; \text{"FL"}; \text{"PS"})
\]

\[
\text{IF}(C5<50 ; \text{"FL"}; \text{IF}(C5\geq85 ; \text{"HD"}; \text{"Between PS and DN"}))
\]
Logical (Boolean) Functions

Logical Functions combine conditions:

- **AND**: Returns TRUE if all its arguments are TRUE; returns FALSE if one or more arguments is FALSE. Example:

  \[
  \text{IF(AND(B2>75; C2>15); "Very Good"; "Room for improvement")}
  \]

- **OR**: Returns TRUE if any argument is TRUE; returns FALSE if all arguments are FALSE. Example:

  \[
  \text{IF(OR(B2>75; C2>15); "Good"; "Must try harder")}
  \]

- **NOT**: Reverses the value of its argument, a logical expression. Example:

  \[
  \text{IF(NOT(C5<50); "PS"; "FL")}
  \]

AND and OR accept up to 30 arguments.
Quiz

- Cell C6 contains the formula as shown below
- What formula will you get if you copy and paste the formula into the cells:

1. A6
2. C4
3. A4

<table>
<thead>
<tr>
<th></th>
<th>A</th>
<th>B</th>
<th>C</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6</td>
<td></td>
<td></td>
<td>$K9+1</td>
</tr>
<tr>
<td>7</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Conditional Formatting

- We can **highlight** cells that meet specific conditions. The formats supplant any original formatting, and are updated whenever cell values change.

**Procedure:**
- Select a cell or a range of cells
- Pick *Format* – *Conditional Formatting*
- Apply up to three conditions and associated formatting styles
- Conditions either
  - compare the current **cell value** with something, or
  - apply a **formula** relative to the current cell
- **Styles** can be selected from a list, or created on the spot (these are remembered indefinitely, so should be given meaningful names)

**Typical usage is to highlight** (with a distinctive background colour):
- Cell values that are unusually high or low (what’s “usual” could be calculated from other data on the sheet)
- Local relationships such as “same as value above”
- Complete rows based on criteria in one column
Conditional Formatting References

• Getting the conditions right can be tricky because one formula is entered to cover the entire range where the formatting applies
  1. Cell values compared to constants – no issues
  2. Cell values compared to absolute references – no issues
  3. Cell values compared to relative locations or condition based on a formula
     - Crucial point: use mixed or relative addressing assuming the current cell is the active cell in the selected range, usually the last one
     - Example: highlight cell values that are greater than the one above

Active cell is A35, so “the cell above” is A34. Use $A34 or A34 but not $A$34 or every highlight will refer to that cell (32)
Conditional Formatting References

- When using a formula, you need to be aware whether to use mixed or relative addressing, that is, which rows or columns are fixed and which have to change with the formula
  - Example from the Sturec_3 sheet: highlight rows 3 to 45 (up to column I) where the Total mark is more than 10 more than the average Total mark
  1. Select A3:I45 (if you work top-down and left-right then I45 is the active cell, but if you work bottom-up it could be A3 or I3)
  2. Assuming I45 is the active cell, enter the formula

\[
\text{"$G45 > \text{AVERAGE}($G$3:$G$45)+10"}
\]

The first reference must be mixed..

Notes: If you point to cells when composing formulas, the full absolute address is used, including the sheet name (usually harmless).

Excel’s conditional formatting includes some useful presets, but entering formulas isn’t as easy as Calc, and references are always to the top left cell of the range.
Unit Conversions

- Calc offers a handy function called **CONVERT_ADD***, that can be used to carry out simple and complex unit conversions from one system to another, including:
  - `convert_add(525; "mi"; "m")` convert constant values
  - `convert_add(B45; "g"; "lbbm")` convert values in formulas
  - `convert_add(1; "gal"; "L")` 1 => unit conversion for use as multiplier
  - `convert_add(1; "atm"; "hPa")` SI prefixes can be applied to SI and some other units. Prefixes are k, M, G, T, …; m, u, n, p,…; c, d, e, h
  - `convert_add(34; "C"; "F")` non-zero intercept (not a multiplier)
  - `convert_add(4.5; "m/s"; "mph")` speed (no km/hr though)
  - `convert_add(1; "ft2"; "ha")` area

- Full table, including which units accept prefixes, is at

* Excel's equivalent is called **convert**, naturally enough, but that name was hijacked ages ago for a Calc function to do internal Euro currency conversions. Excel doesn’t know about areas or volume measures like m³, and certainly doesn’t know how much beer a Schooner holds (Calc does, provided you spell it properly and don’t spill any 😊).
Useful Functions and Features

Text functions

- Calc offers a wide range of functions to handle text, including
  - `len(text)`, number of characters
  - `upper(text)`, converts text to uppercase, also `lower` and `proper`.
  - `concatenate(text1, text2, ...)`, can also use the `&` operator
  - `trim(text)`, removes surrounding spaces from text
  - `mid(text, start, n)`, returns up to n characters from text, given the starting position (counting from 1), also `left` and `right` without `start`
  - `substitute(text, searchtext, replacement, times)`, find-and-replace

Spreadsheet lookup functions (week 3)

- search columns or rows for matching value, return related data, a bit like a primitive database

Financial functions (week 3)
Counting number of heart beats

# heart beats = # tall peaks
Peak or not?

- Can you come out with a logical test to determine whether a point is a peak or not?

<table>
<thead>
<tr>
<th>Time</th>
<th>Voltage</th>
</tr>
</thead>
<tbody>
<tr>
<td>20.30</td>
<td>1.91</td>
</tr>
<tr>
<td>Point A</td>
<td></td>
</tr>
<tr>
<td>20.40</td>
<td>2.02</td>
</tr>
<tr>
<td>20.50</td>
<td>3.57</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Time</th>
<th>Voltage</th>
</tr>
</thead>
<tbody>
<tr>
<td>22.10</td>
<td>4.11</td>
</tr>
<tr>
<td>Point B</td>
<td></td>
</tr>
<tr>
<td>22.20</td>
<td>2.87</td>
</tr>
<tr>
<td>22.30</td>
<td>2.30</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Time</th>
<th>Voltage</th>
</tr>
</thead>
<tbody>
<tr>
<td>22.80</td>
<td>3.20</td>
</tr>
<tr>
<td>Point C</td>
<td></td>
</tr>
<tr>
<td>22.90</td>
<td>4.03</td>
</tr>
<tr>
<td>23.00</td>
<td>2.72</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Time</th>
<th>Voltage</th>
</tr>
</thead>
<tbody>
<tr>
<td>23.30</td>
<td>1.97</td>
</tr>
<tr>
<td>Point D</td>
<td></td>
</tr>
<tr>
<td>23.40</td>
<td>1.90</td>
</tr>
<tr>
<td>23.50</td>
<td>2.10</td>
</tr>
</tbody>
</table>
Counting number of heart beats

• Now you know how to determine whether a point is a peak

• However, not all the peaks are tall peaks

• We only want to count those peaks whose voltage exceeds a threshold
  – This means we need an additional logical condition
How computers will change the world

• Computers have changed and will change engineering
  – That’s why computing is important for engineers
• The world will become more automated in the future and computers will play a key role
• To round up Week 1, here is a vision for the future.
• Video at: https://www.google.com/selfdrivingcar/