<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 Organisation</td>
<td>1</td>
</tr>
<tr>
<td>1.1 Color tabs</td>
<td>1</td>
</tr>
<tr>
<td>1.2 Hide tabs in dashboard</td>
<td>1</td>
</tr>
<tr>
<td>1.3 Organise Datafields</td>
<td>2</td>
</tr>
<tr>
<td>2 Types of Visuals</td>
<td>3</td>
</tr>
<tr>
<td>3 Axis</td>
<td>5</td>
</tr>
<tr>
<td>3.1 Multiple Measures</td>
<td>5</td>
</tr>
<tr>
<td>3.2 Time Series (w/o) Date</td>
<td>5</td>
</tr>
<tr>
<td>4 Filters</td>
<td>7</td>
</tr>
<tr>
<td>4.1 Filter Last Date</td>
<td>7</td>
</tr>
<tr>
<td>5 Field Calculations</td>
<td>9</td>
</tr>
<tr>
<td>5.1 Level of Detail (LOD)</td>
<td>9</td>
</tr>
<tr>
<td>6 SQL</td>
<td>11</td>
</tr>
<tr>
<td>6.1 Datetime Slicing</td>
<td>11</td>
</tr>
<tr>
<td>6.2 Restrict Datetime Range</td>
<td>11</td>
</tr>
<tr>
<td>6.3 Gap Fill</td>
<td>12</td>
</tr>
<tr>
<td>6.4 Parameter with (ALL) option</td>
<td>13</td>
</tr>
<tr>
<td>7 Dashboard Layout</td>
<td>15</td>
</tr>
<tr>
<td>7.1 Device Preview</td>
<td>15</td>
</tr>
<tr>
<td>7.2 Dashboard Background</td>
<td>15</td>
</tr>
<tr>
<td>8 Custom Stuff</td>
<td>17</td>
</tr>
<tr>
<td>8.1 Color Palettes</td>
<td>17</td>
</tr>
<tr>
<td>8.2 Shapes</td>
<td>17</td>
</tr>
<tr>
<td>9 Drivers</td>
<td>19</td>
</tr>
<tr>
<td>9.1 Impala</td>
<td>19</td>
</tr>
</tbody>
</table>
Color tabs

Color coding worksheets makes it easier to differentiate which group them belong to.

Hide tabs in dashboard

If you have too many worksheets, you can hide them by right clicking a dashboard and hide all associated worksheets. Unhide all of them with the same function.

To show each of them individually to worksheet view, there are two options. Note that they will be hidden when you go to another tab. You will need to right-click that worksheet tab and unhide it for it to stay.
Organise Datafields

For proper organisation, it is better to create folders and drop datafields into folder categories. We can search for specific fields using the search box.
CHAPTER 2

Types of Visuals
Multiple Measures

Time Series (w/o) Date

In SQL or Field Calculation, set a datetime field to time with date as a constant.

Add field to visual row/column. Change format to just hh:nn or h:nn AMPM to prevent the date constant from showing.
CHAPTER 4

Filters

Filter Last Date

This trick is sourced from http://kb.tableau.com/articles/howto/filter-on-last-date-in-the-data-source. It is achieved from the filter pop-up, top tab.
Chapter 5

Field Calculations

Level of Detail (LOD)
For production development, it is preferred to do most of the ETL and especially aggregation in the database so that data being fed to Tableau is much smaller. This is a key performance requirement since the database engine is usually more powerful.

**Datetime Slicing**

`#-- Show Only Time at Axis
#-- Set date as constant so that Tableau can group by time only
SELECT to_char(DATE_TIME, '1900-01-01 HH24:MI')::timestamp
FROM table_nm`

`#-- Time Slice
#-- change 900 to appropriate seconds, e.g., 15 mins == 15 * 60 = 900
#-- convert timestamp to epoch, divide by interval desired in minutes then round to get the desired interval
SELECT to_timestamp(floor((extract('epoch' from starttime) / 900 )) * 900)
FROM tablename`

**Restrict Datetime Range**

`#-- Filter Date Range
WHERE trunc(DATE_TIME) BETWEEN :strt_l_tme and :end_l_tme;

#-- Filter Time Range
WHERE (to_char(DATE_TIME,'hh24:mi') between :stme1 and :etme1)`
#--Filter Fixed Time Range
where transactiontimestamp between now()::timestamp - interval '24 hour'
       and now()::timestamp

#--Filter Latest Data by DateTime
where transactiontimestamp = (select max(transactiontimestamp) from joins limit 1)

Gap Fill

Gap Filling is very important to fulfil the aesthetics of a visualisation. Lets say a barchart is supposed to have 1 dimension with 10 attributes to show. But the dataset only contain 5 at this time. The barchart will only show 5 bars. This is bad for aesthetics as the bar width will resize whenever a new row show a new attribute of that dimension.

Fig. 6.1: Other attributes will not show as none of the rows have those data

To prevent that, we can have a union in sql to add all possible attributes in.

    select col1, col2, dimension from table1
    union
    select null as col1,
           null as col2,
           dimension
    from mapping_table

Fig. 6.2: Other attributes will show now.

Another time of gap fill occurs when we have a visualisation with more than 1 dimension. The previous code will not do as it will produce this. A example is a heatmap shown below.

A cross-join for the new table for union in SQL is necessary for this case to fill every single gap for both dimensions. Now it is perfect!

    #--generate_series() is exclusive for postgresql
    select col1, col2, dimension1, date_dimension2 from table1
    from table1
    where date_dimension2 between '2007-02-01'::timestamp and '2008-04-01'::timestamp
Parameter with (ALL) option

One of the disadvantages of Parameters to Quick Filters among other things is that it lacks a select ALL option. However, we can manually code for this in SQL.

```sql
-- with (All) option
--either define a table with possible and (All) attributes, and define in where clause
with filter_week AS (select 'Weekend' AS weektype, 'Weekend' AS param
    union select 'Weekday' AS weektype, 'Weekday' AS param
    union select 'Weekday' AS weektype, '(All)' AS param
    union select 'Weekend' AS weektype, '(All)' AS param)
SELECT * FROM DBO.table_nm a
WHERE WEEKTYPE IN (SELECT weektype FROM filter_week WHERE param=:week);
```
-- or just entirely in the WHERE clause as below
WHERE (((time_belt='Morning' OR time_belt='Day' OR time_belt='Evening') AND
     (time_belt='(All)') OR (time_belt=:timebelt))}
CHAPTER 7

Dashboard Layout

Device Preview

Dashboard Background
CHAPTER 8

Custom Stuff

Color Palettes

New color palettes can be added at the Documents\My Tableau Repository\Preferences.tps file. You can add new palettes manually by editing the file, or replace it with another tps that someone build. Note to change the name back to Preferences.tps.

An excellent list of palettes is from the famous Color Brewer. Someone has created it for Tableau, which can be downloaded here.

Shapes

New shapes/symbols can be added to Tableau by creating a new folder or adding into existing folders within Documents\My Tableau Repository\Shapes.
Impala

Impala is a massively parallel processing SQL query engine for Hadoop. In short, it is very fast when running queries, with significant performances over Hive.

However, connection to Tableau Desktop or Tableau Server might not be as straightforward, especially in Mac or Linux.

For a start:

• Take a look at https://www.tableau.com/support/drivers
• Go to https://www.cloudera.com/documentation/other/connectors.html
• Select OS & driver connection type & version
• Driver will be downloaded
• Also download documentation at same page or use the link below
• Follow the documentation for driver setup, main thing is to configure the DSN (Data Source Name)

Windows

First, go to Windows ODBC manager. Go to Windows > Search > 64-bit ODBC Manager.

Go to Impala IDE, HUE and check the database name to connect.

Note that Tableau’s Cloudera Hadoop ODBC driver does not work, god knows why. Click on Other Databases (ODBC) and enter the details for DSN name, Server IP, & Port (51052)
MacOS

- cd /opt/cloudera/impalaodbc/Setup
- copy odbc.ini & odbcinst.ini to user directory
- cp odbc.ini ~
- cp odbcinst.ini ~
- open odbc.ini
  - Change data source name (DNS) to your liking
  - add HOST=[your_ip]
  - add PORT=[21050]
- change both files to hidden
  - mv odbc.ini .odbd.ini
  - mv odbcinst.ini .odbcinst.ini
- Locate where iodbc driver manager libraries are located. Use sudo find / -name "*iodbc* to search
- Then add the following environment variables to the bash profile nano .bash_profile

```bash
export DYLD_LIBRARY_PATH=$DYLD_LIBRARY_PATH:~/Volumes/Macintosh HD/usr/lib/
export ODBCINI=~/.odbc.ini
```

export ODBCINSTINI=~/.odbcinst.ini

Fig. 9.1: HUE IDE

Fig. 9.2: Add the host, port & database
export CLOUDERAIMPALAODBCINI=~/.cloudera.impalaodbc.ini

**LINUX (UBUNTU)**

- Same as Mac but for .cloudera.impalaodbc.ini, change the default iodbc manager name to what is indicated in system, i.e., `libiodbc.so` to `libiodbc.so.2`
- Locate where iodbc driver manager libraries are located. Use `sudo find / -name "*iodbc*` to search
- Go to `nano ~/.bashrc` and enter the following environmental variables

```
export DYLD_LIBRARY_PATH=$DYLD_LIBRARY_PATH:/usr/lib/x86_64-linux-gnu/
export ODBCINI=~/.odbc.ini
```

- `export ODBCINSTINI=~/.odbcinst.ini`
- `export CLOUDERAIMPALAODBCINI=~/.cloudera.impalaodbc.ini`
  - Use `iodbctest` `dsn=Tableau Impala Connection` to test connection to Impala
  - Enter some SQL commands to test

**Tableau Desktop**

- In Tableau Desktop, go to Other Databases (ODBC)
- At DSN > select the DSN name you created
- Click Connect
- Enter Server, Port (21050) & database (default)
- Click Sign In