



# Melbourne Datathon 2017



Welcome to the 2017 Melbourne Datathon!

There are 2 components to this year's event, the *Insights* competition and the *Kaggle* competition

## Insights Competition

Your task is to convince your “bosses” that they need to invest in analytics. Do this by investigating the data and putting together a presentation that will sell its value.

The insights competition will run for 20 days until 3pm on Wed 3<sup>rd</sup> May, at which time all submissions must have been made. Submissions must be in pdf format with the filename that of the team. On the first slide, please include the team name and team members with email addresses. Email your entries to [Submissions2017@datasciencemelbourne.com](mailto:Submissions2017@datasciencemelbourne.com). You should get an auto confirmation response but we will post all the submitted entries later that evening on the datathon website, so contact us if things went astray ([www.datasciencemelbourne.com/datathon](http://www.datasciencemelbourne.com/datathon)). Note: the attachment size limit is 20mb.

The top 5 teams will be announced by 12 noon on Thursday and asked to pitch their findings to the judging panel from 5:30pm on Friday 5<sup>th</sup> May when the overall winners will be determined. Select entrants will subsequently be invited to present in a special session at the Wombat MeDaScIn conference on 2<sup>nd</sup> June during Melbourne Data Science Week.

Some ideas for investigation might be;

- Are there any data quality issues?
- Are there any demographic patterns in illnesses?
- Which pharmacies have best patient loyalty; is loyalty age skewed?
- Distribution of packet sizes (are larger pack sizes favoured, why? Does this change by geographical location?)
- Are generic medicines favoured by any geographical location, age or gender?
- Any explanation of the seasonal demand patterns?
- Any specific seasonal patterns for particular drugs?
- Any suspicious behavior in the data?
- Can patients be clustered by drug usage patterns?
- Any insights about where the trajectory of the health of Australia is heading?

These are just suggestions but you will be looking at this data in more detail than it has ever been looked at before, so you tell us where the gold is.



# Melbourne Datathon 2017



## Kaggle Competition

The Kaggle competition is now live and runs for 45 days until noon on Tue 30<sup>th</sup> May. The top 5 teams on the leaderboard as of 12 noon on Sun 28<sup>th</sup> May will each receive 2 complimentary tickets to the Wombat MeDaScIn conference on the 2<sup>nd</sup> June, where the winners will be formally announced.

Teams for the Kaggle competition do not need to be the same as for the Insights competition. If you intend to apply for an internship, then it is recommended to enter the Kaggle competition as an individual.

Hope to see you soon on the leaderboard...

<https://inclass.kaggle.com/c/dsm2017>

## Internships

If you are considering applying for an internship, then a submission to the insights competition must first have been made – this can be as an individual or as part of a team.

Your application should be your CV and then insights deck combined into a single pdf document named FIRSTNAME\_LASTNAME.pdf. The first page should include the following information:

1. Name
2. Email address
3. Phone number
4. Kaggle name (so we can locate you on the leaderboard)
5. Insights team name (also indicate if you were placed in the top 5)
6. Any working restrictions attached to your visa and visa expiration date (if not applicable just say 'no working restrictions').
7. Tell us what you are looking for (ie are you still at Uni or are you looking for a permanent role) and when you are available to start
8. If you would like to be considered for any other roles that come up in the future then please indicate so.

Please email your applications to [Internships2017@datasciencemelbourne.com](mailto:Internships2017@datasciencemelbourne.com) by 5pm on Monday May 8<sup>th</sup> (20mb limit). All CVs will then be forwarded directly to the companies involved who will contact prospective interns directly. Unfortunately this is the last involvement we will have, so we won't be able to let you know how your application is progressing.



# Melbourne Datathon 2017



## Social Media

Please use #MelbourneDatathon and follow @DataScienceMelb on twitter and <https://www.facebook.com/melbournedatathon/>.

There will be a spot prize of a free conference ticket for the first person to tweet how many rows of data there are and other spot prizes throughout the event.

## Thanks

This free event would not be possible without the direct support of the following organisations and the continued support we receive from the Melbourne analytics community throughout the year.

