Measuring Impact Quantitatively

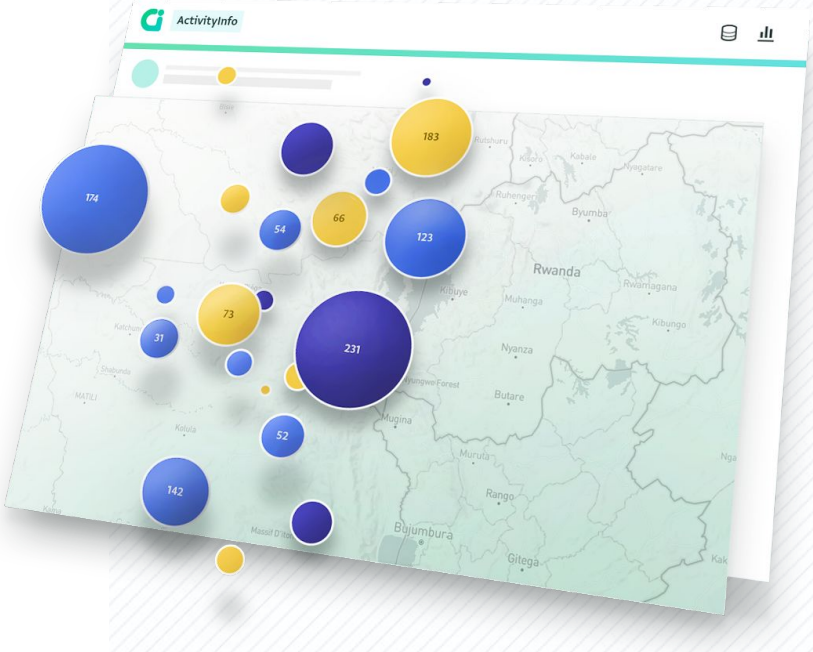
May 10th, 2022

ActivityInfo

Presented by the ActivityInfo Team

Monitoring & Evaluation Software

- Track activities, outcomes
- Beneficiary management
- Surveys
- Work offline / online



Activity Info

Outline

1. Quick review
2. Statistical significance
3. Effect size
4. Questions!

Quick review

Quantitative Impact Evaluation

Causal impact

Outcome Y *without* the program

$$\Delta = (Y \mid P = 1) - (Y \mid P = 0)$$

Outcome Y *with* the program

Activity Info

Key points from Part I

- Why conduct a quantitative impact evaluation?
- When would you not conduct a quantitative impact evaluation?
- Types of measurements
- Sources of measurement error
- Reliability: Cronbach’s alpha
- Cognitive interviewing, a tool for improving questionnaires

Activity Info

Key point from Part II

- Fundamental problem of causal inference
- Identify four strategies for “counterfeit counterfactuals”
- Identify risks of before-and-after comparisons

Activity Info

Statistical significance

Null hypothesis

The hypothesis that our program has had **ZERO** impact.

Activity Info

Null hypothesis

Causal impact

Outcome Y *without* the program

$$0 = (Y \mid P = 1) - (Y \mid P = 0)$$

Outcome Y *with* the program

Activity Info

Significance testing

Can we disprove the null hypothesis?

Significance testing

IF the null hypothesis was true, what are the chances of sampling the data we sampled?

Hypothesis testing: simple example

Hypothesis: the coin is fair --
the odds of flipping a coin are 50/50

Experiment: 3 heads, 0 tails

What are the chances if the coin is fair?

Activity Info

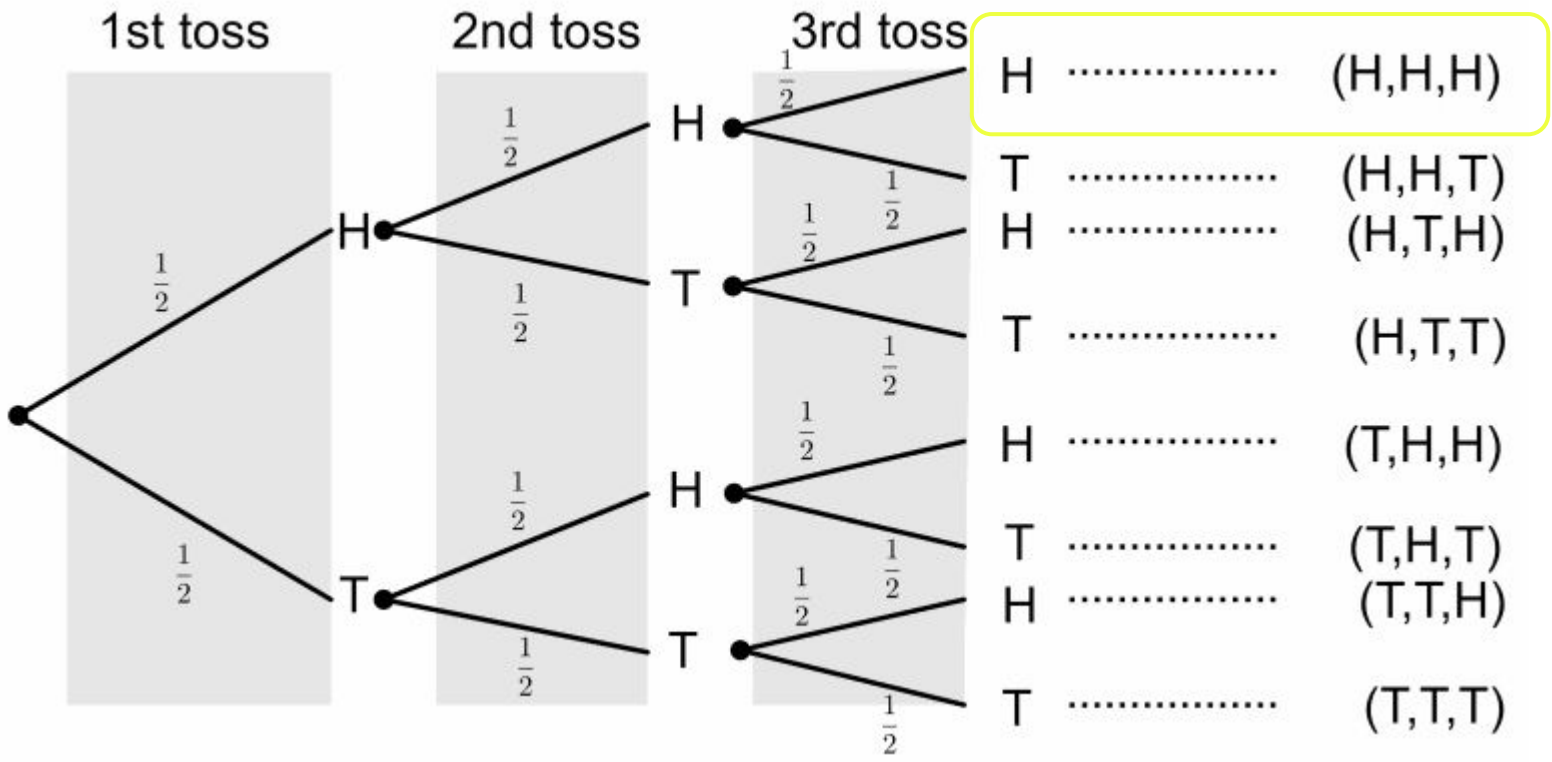


Image source: https://www.storyofmathematics.com/coin-flip-probability/

Hypothesis testing: simple example

Hypothesis: the coin is fair --
the odds of flipping a coin are 50/50

Experiment: 3 heads, 0 tails

What are the chances if the coin is fair?
12.5%, $p = 0.125$

Can we reject the hypothesis?

No, this is not *so* unusual, even if the coin was fair.

Activity Info

Hypothesis testing: simple example

Hypothesis: the coin is fair --
the odds of flipping a coin are 50/50

Experiment: 8 heads, 2 tails

What are the chances if the coin is fair?
4% probability, $p = 0.04$

Can we reject the hypothesis?

Borderline. This is somewhat unusual, but it would still happen ~ 1 out of 20 times.

Activity Info

Hypothesis testing: simple example

Hypothesis: the coin is fair --
the odds of flipping a coin are 50/50

Experiment: 19 heads, 1 tails

What are the chances if the coin is fair?
0.001907% probability, $p < 0.0001$

Can we reject the hypothesis?

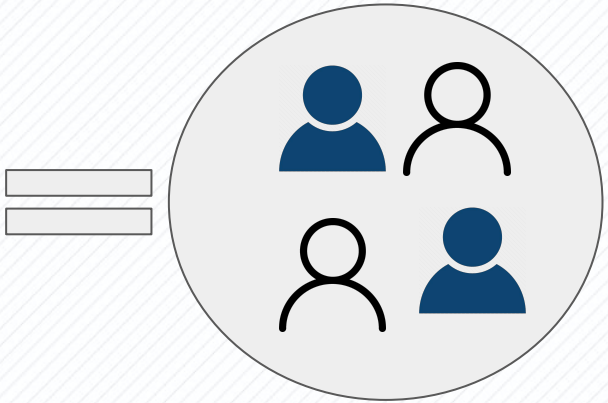
Yes! If the coin was fair the chances of this happening are freakishly low.

Hypothesis testing: impact

$(Y \mid P = 1) = 60\$/month$

$(Y \mid P = 0) = 50\$/month$

+10$

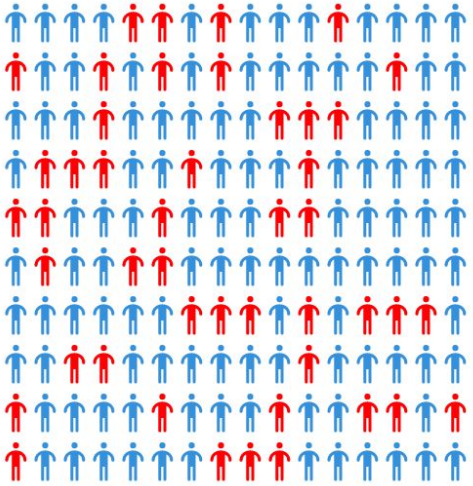


Randomly to receive training

Control group who received no training

Activity Info

Hypothesis testing



(Null) Hypothesis: our training program has zero impact on incomes

Experiment: Sample 50 people from each group, difference is +10 USD.

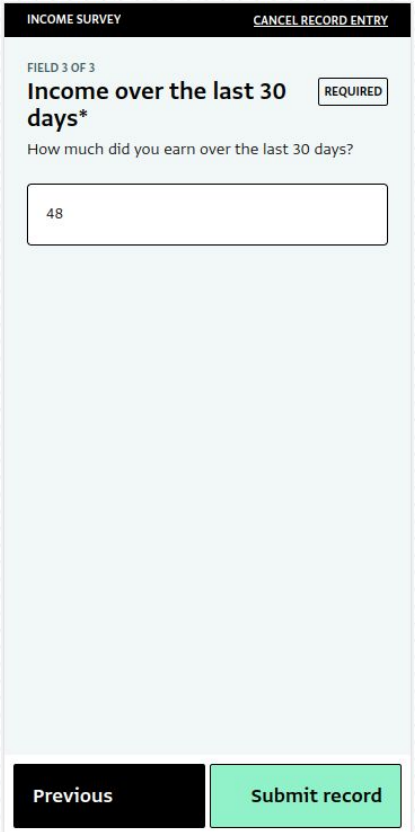
What are the chances if our training program has no impact?

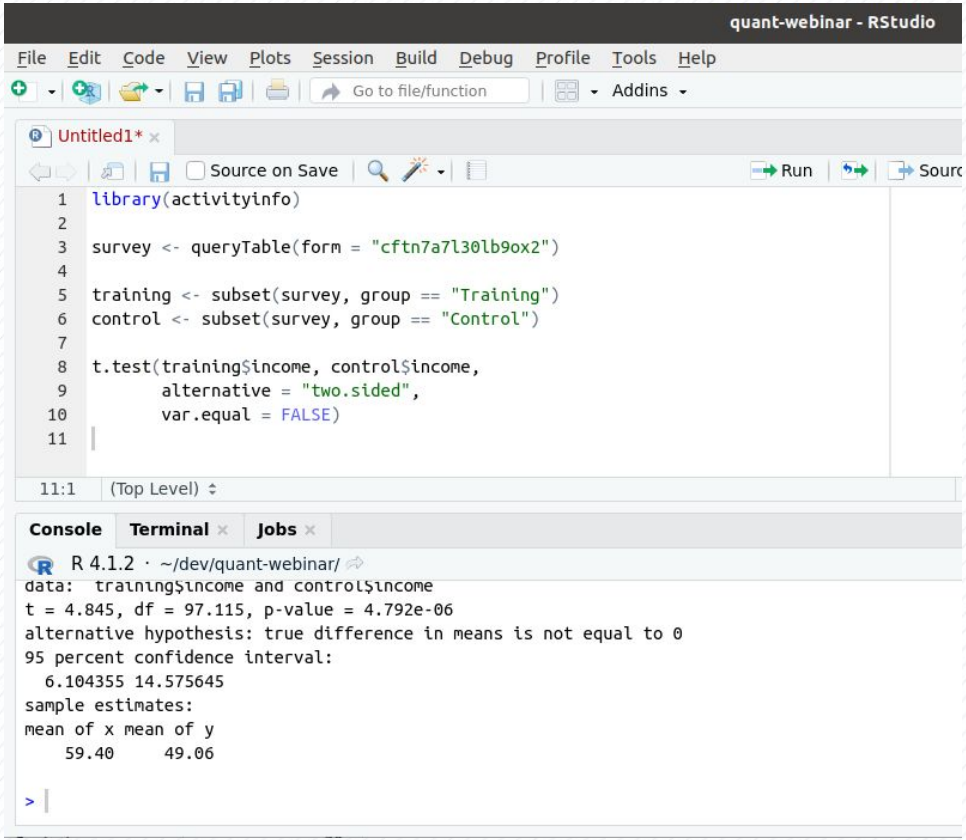
Activity Info

Calculating probability of mean difference

Student’s t-test calculates the probability of getting this much difference in averages, if the null hypothesis were true.

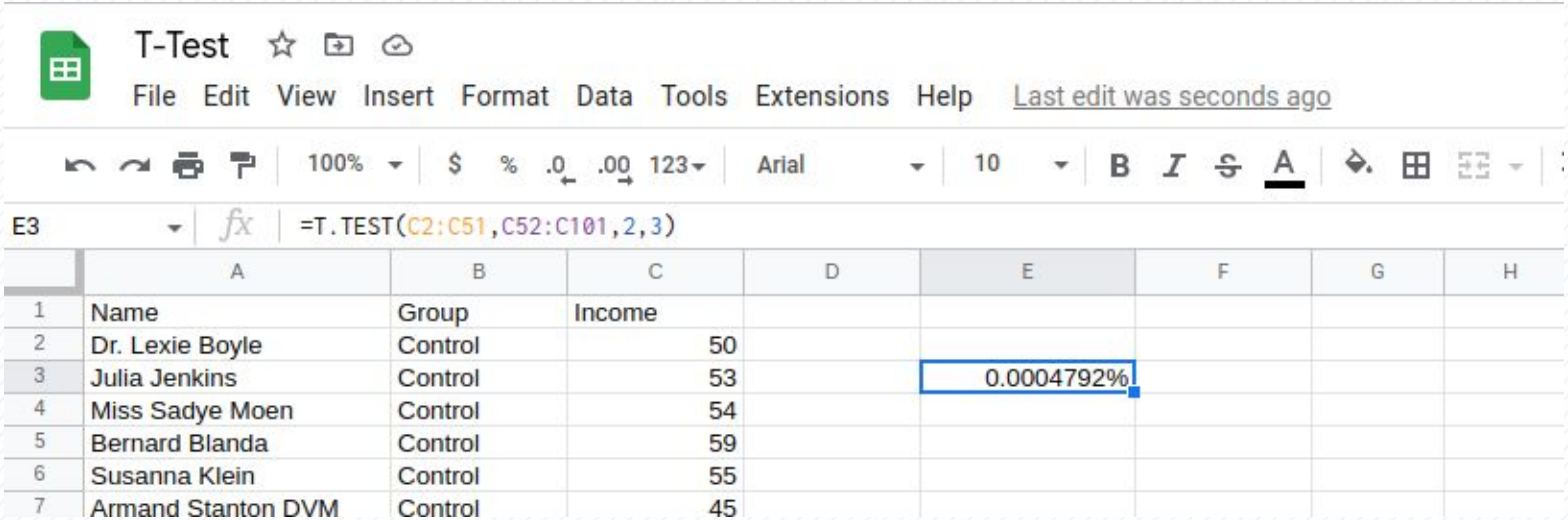
Data collection > Analysis





Activity Info

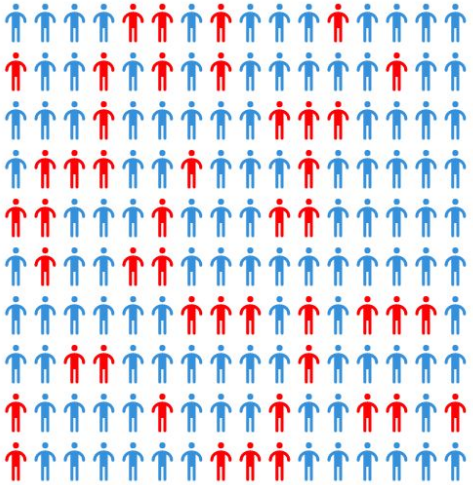
t-Tests with Excel



	A	B	C	D	E	F	G	H
1	Name	Group	Income					
2	Dr. Lexie Boyle	Control	50					
3	Julia Jenkins	Control	53		0.0004792%			
4	Miss Sadye Moen	Control	54					
5	Bernard Blanda	Control	59					
6	Susanna Klein	Control	55					
7	Armand Stanton DVM	Control	45					

Activity Info

Hypothesis testing



(Null) Hypothesis: our training program has zero impact on incomes

Experiment: Sample 50 people from each group, difference is +10 USD.

What are the chances if our training program has no impact? 0.0004792%

Can we reject the hypothesis?

Yes! If our program had no impact, there would be practically no way of getting these results by chance.

Activity Info

Hypothesis standards

$p < 0.05$ * “statistically significant”

$p < 0.01$ ** “highly statistically significant”

$p < 0.001$ ***

Activity Info

Hypothesis testing - pitfalls

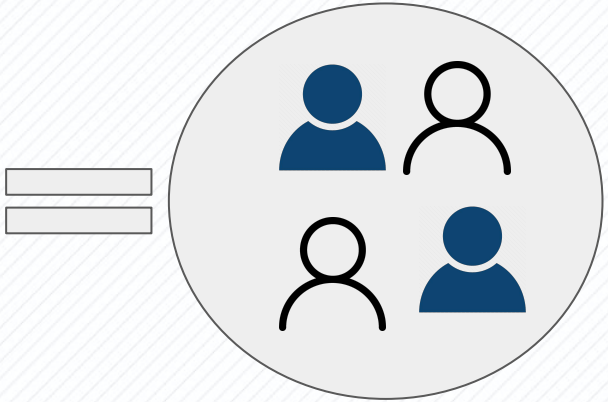
“Not zero” is a very low standard is a very low standard for impact!

Hypothesis testing: small effects

(Y | P = 1) = 51$/month

(Y | P = 0) = 50$/month

+1$



Randomly to receive training

Control group who received no training

Activity Info

Miniscule differences can still be “significant” with enough data!

Impact of program: +1 USD

Sample size 50, p = 0.519

Sample size 100, p = 0.236

Sample size 500, p = 0.014*

Sample size 1000, p = 0.000***

Activity Info

Effect size

Effect size

Statistics which help understand the **size** of our impact, not just whether the impact is non-zero.

Examples of effect size statistics

- Difference in means (averages)
- Cohen’s d
- Odds ratios
- Eta-squared

Revisiting our small size example

Impact of program: +1 USD

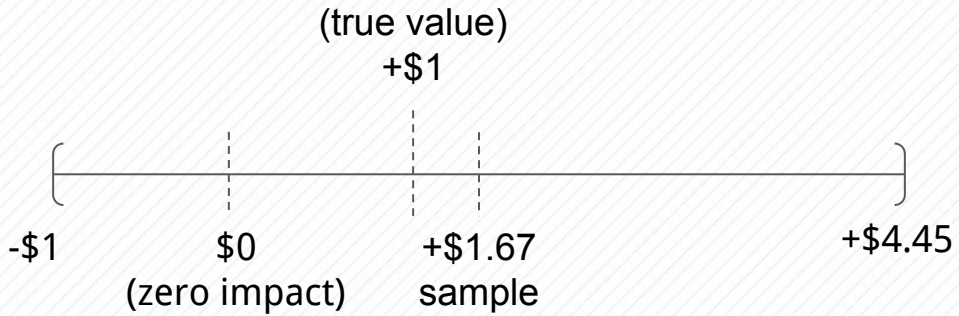
Sample size 50, $p = 0.519$, [-2.42 - 4.75]

Sample size 100, $p = 0.236$, [-1.11 - 4.45]

Sample size 500, $p = 0.014$, [0.34 - 2.96]

Sample size 1000, $p = 0.000$, [1.12 - 2.95]

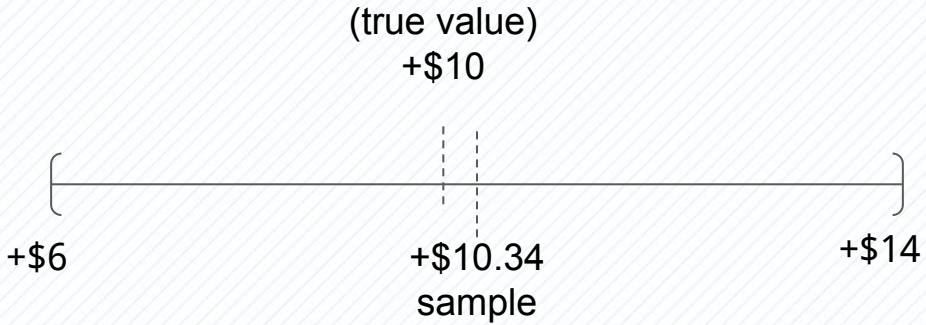
Confidence intervals



Sample size 100, $p = 0.236$, [-1.11 - 4.45]

Activity Info

Confidence intervals



Sample size 50, $p < 0.001$, [6.10 - 14.57]

Activity Info

Cohen’s d

- *Standardized* mean difference
- Useful for values without easily interpretable units

$$d = \frac{\bar{x}_1 - \bar{x}_2}{s}.$$

Activity Info

Interpreting Cohen's d

Roughly speaking, how much variation is explained

0.5494623 1.3885407

Effect size	*d*
Very small	0.01
Small	0.20
Medium	0.50
Large	0.80
Very large	1.20
Huge	2.0

Note on complex samples

If your sample uses clusters or stratification, you **cannot** use a simple t-Test. See the R survey package:

https://cran.r-project.org/web/packages/survey/index.html

```
design <- svydesign(id=~dnum+snum, data=training)

svyttest(enroll~income, dclus2)
```

Communicating about effect size

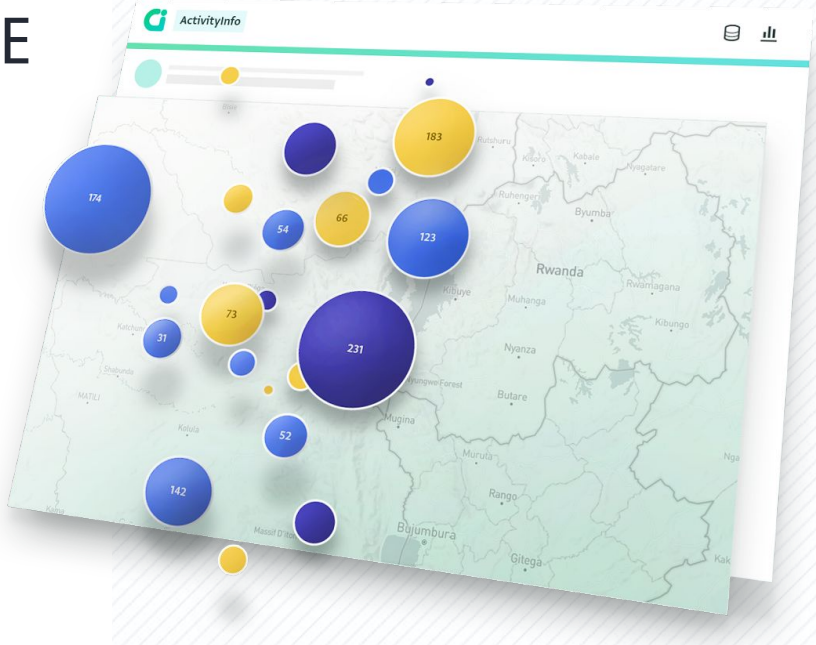
Try ActivityInfo!

The most complete database system for M&E available.

[Try a template]

Demo Webinar:

May 31st, 14:00 CEST



Activity Info