



GAM130 Development Practice  
**Game Analytics**



# What is analytics?



# A rhetorical question

# A rhetorical question

- ▶ Suppose you could put your game in front of **thousands of players**

# A rhetorical question

- ▶ Suppose you could put your game in front of **thousands of players**
- ▶ You can alter your game's code to **measure** things about how players play and **report** them back to you

# A rhetorical question

- ▶ Suppose you could put your game in front of **thousands of players**
- ▶ You can alter your game's code to **measure** things about how players play and **report** them back to you
- ▶ What **questions** could you answer about your game?

# A rhetorical question

- ▶ Suppose you could put your game in front of **thousands of players**
- ▶ You can alter your game's code to **measure** things about how players play and **report** them back to you
- ▶ What **questions** could you answer about your game?
- ▶ How could these answers make your game more **enjoyable** and/or **profitable**?

# Analytics

# Analytics

- ▶ Analytics is the analysis of **data** to discover **patterns**

# Analytics

- ▶ Analytics is the analysis of **data** to discover **patterns**
- ▶ Related terms: data mining, data science, big data,

...

# Why analytics?

# Why analytics?

- ▶ Most of our gaming devices nowadays have **always-on internet connections**

# Why analytics?

- ▶ Most of our gaming devices nowadays have **always-on internet connections**
- ▶ Developers can easily push **updates** to released games

# Why analytics?

- ▶ Most of our gaming devices nowadays have **always-on internet connections**
- ▶ Developers can easily push **updates** to released games
- ▶ Games can easily **collect data** and send it back to developers

# Why analytics?

- ▶ Most of our gaming devices nowadays have **always-on internet connections**
- ▶ Developers can easily push **updates** to released games
- ▶ Games can easily **collect data** and send it back to developers
- ▶ Bandwidth, storage, compute time are **cheap**

# Why analytics?

- ▶ Most of our gaming devices nowadays have **always-on internet connections**
- ▶ Developers can easily push **updates** to released games
- ▶ Games can easily **collect data** and send it back to developers
- ▶ Bandwidth, storage, compute time are **cheap**
- ▶ Analytics can take the **guesswork** out of many game design and business decisions

# Feedback loops

# Feedback loops

- ▶ Game development is a **feedback loop** of **playtesting** and **iteration**

# Feedback loops

- ▶ Game development is a **feedback loop** of **playtesting** and **iteration**
- ▶ Gathering and interpreting informal feedback from playtesters is **hard**

# Feedback loops

- ▶ Game development is a **feedback loop** of **playtesting** and **iteration**
- ▶ Gathering and interpreting informal feedback from playtesters is **hard**
- ▶ ... and even harder when your “playtesters” are your **customers**

# Feedback loops

- ▶ Game development is a **feedback loop** of **playtesting** and **iteration**
- ▶ Gathering and interpreting informal feedback from playtesters is **hard**
- ▶ ... and even harder when your “playtesters” are your **customers**
- ▶ Analytics aims to turn this process into more of a **science**

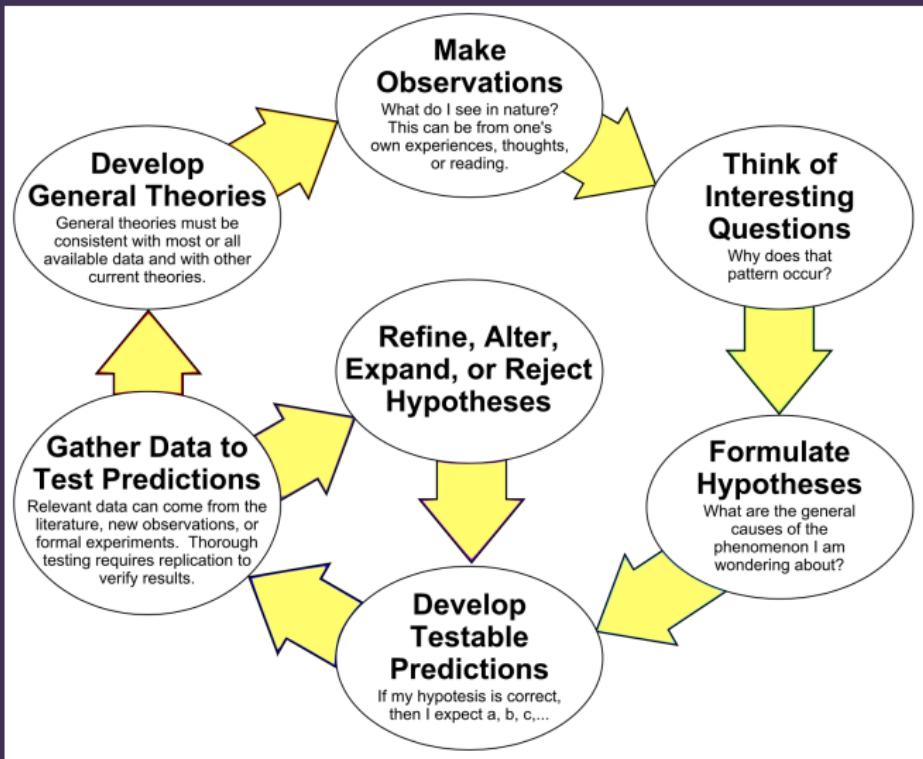
# Feedback loops

- ▶ Game development is a **feedback loop** of **playtesting** and **iteration**
- ▶ Gathering and interpreting informal feedback from playtesters is **hard**
- ▶ ... and even harder when your “playtesters” are your **customers**
- ▶ Analytics aims to turn this process into more of a **science**
- ▶ The **build-measure-learn** cycle

# Analysing data



# The scientific method



# Statistical significance

# Statistical significance

Is a D6 biased if...

# Statistical significance

Is a D6 biased if...

- ▶ You roll it once and it comes up a six?

# Statistical significance

Is a D6 biased if...

- ▶ You roll it once and it comes up a six?
- ▶ You roll it 3 times and it comes up a six twice?

# Statistical significance

Is a D6 biased if...

- ▶ You roll it once and it comes up a six?
- ▶ You roll it 3 times and it comes up a six twice?
- ▶ You roll it 60 times and it comes up a six 59 times?

# Statistical significance

Is a D6 biased if...

- ▶ You roll it once and it comes up a six?
- ▶ You roll it 3 times and it comes up a six twice?
- ▶ You roll it 60 times and it comes up a six 59 times?
- ▶ You roll it 60 times and it comes up a six 11 times?

# Statistical significance

Is a D6 biased if...

- ▶ You roll it once and it comes up a six?
- ▶ You roll it 3 times and it comes up a six twice?
- ▶ You roll it 60 times and it comes up a six 59 times?
- ▶ You roll it 60 times and it comes up a six 11 times?
- ▶ You roll it 600 times and it comes up a six 110 times?

# Statistical significance

Is a D6 biased if...

- ▶ You roll it once and it comes up a six?
- ▶ You roll it 3 times and it comes up a six twice?
- ▶ You roll it 60 times and it comes up a six 59 times?
- ▶ You roll it 60 times and it comes up a six 11 times?
- ▶ You roll it 600 times and it comes up a six 110 times?
- ▶ You roll it 6,000,000 times and it comes up a six 1,100,000 times?

# Statistical significance

# Statistical significance

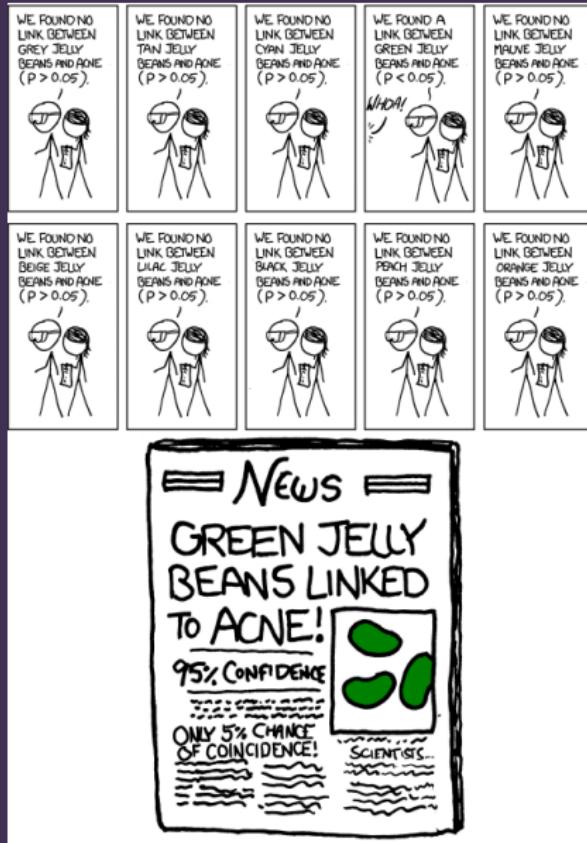
- ▶ Every statistical result has a non-zero probability of being a **coincidence**

# Statistical significance

- ▶ Every statistical result has a non-zero probability of being a **coincidence**
- ▶ Know your **confidence intervals**

# Statistical significance

- ▶ Every statistical result has a non-zero probability of being a **coincidence**
- ▶ Know your **confidence intervals**
- ▶ Beware of *p-value fishing*



# Statistical significance vs effect size

# Statistical significance vs effect size

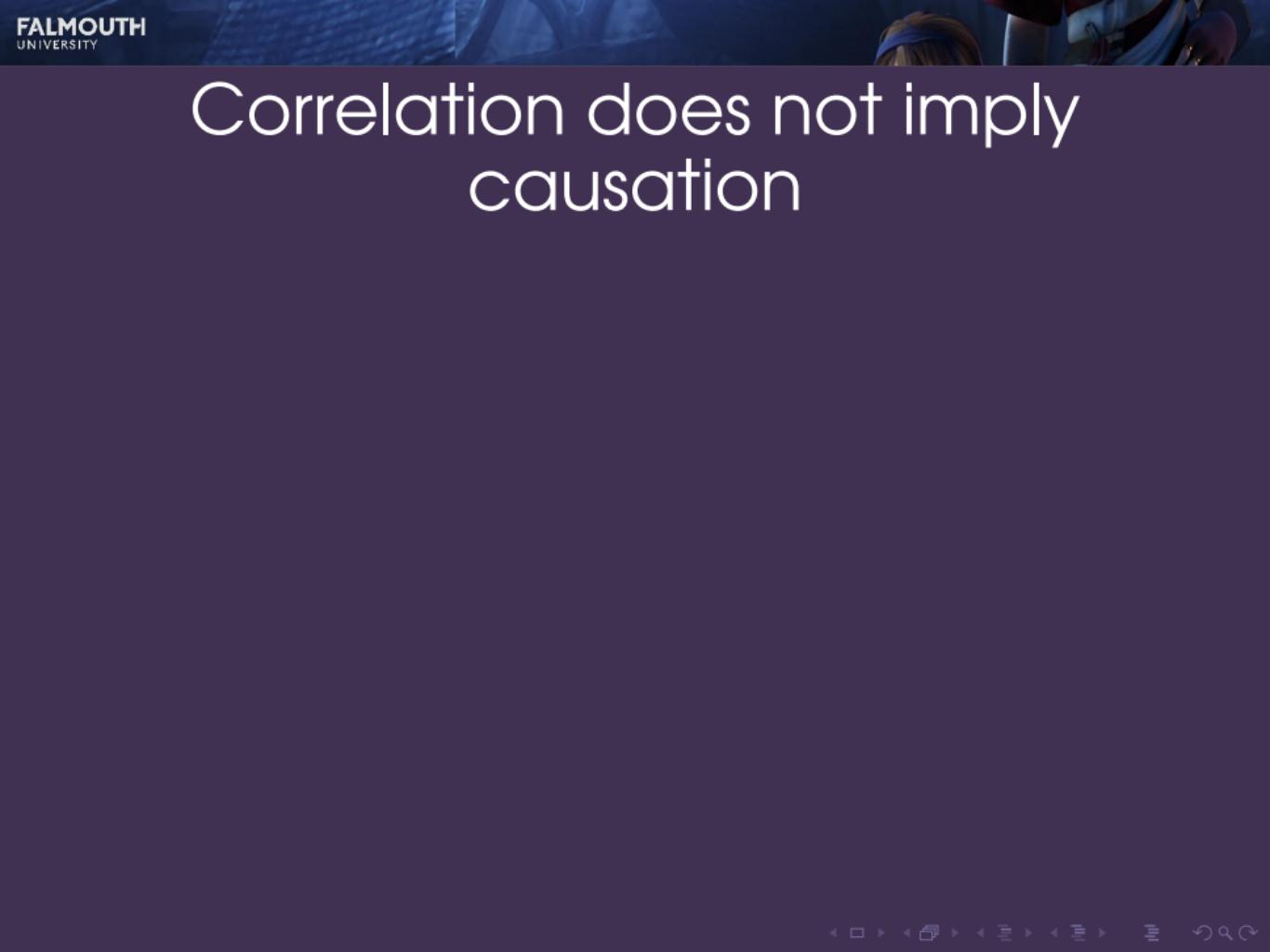
- High statistical significance does not always mean large **effect size**

# Statistical significance vs effect size

- ▶ High statistical significance does not always mean large **effect size**
- ▶ E.g. red team wins 5,010,000 matches out of 10,000,000

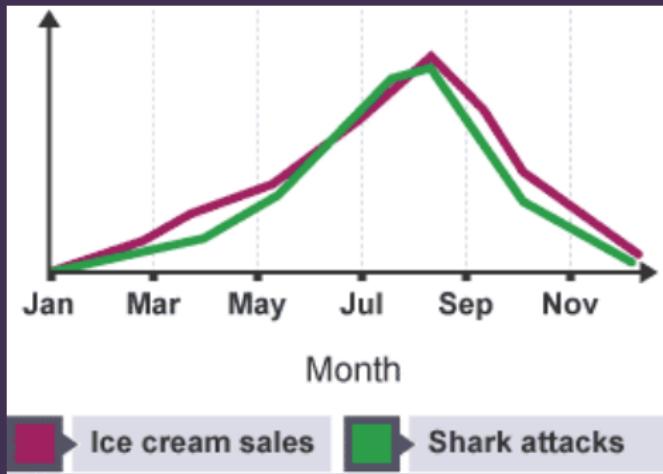
# Statistical significance vs effect size

- ▶ High statistical significance does not always mean large **effect size**
- ▶ E.g. red team wins 5,010,000 matches out of 10,000,000
- ▶ This is statistically significant, but only a 0.1% effect size

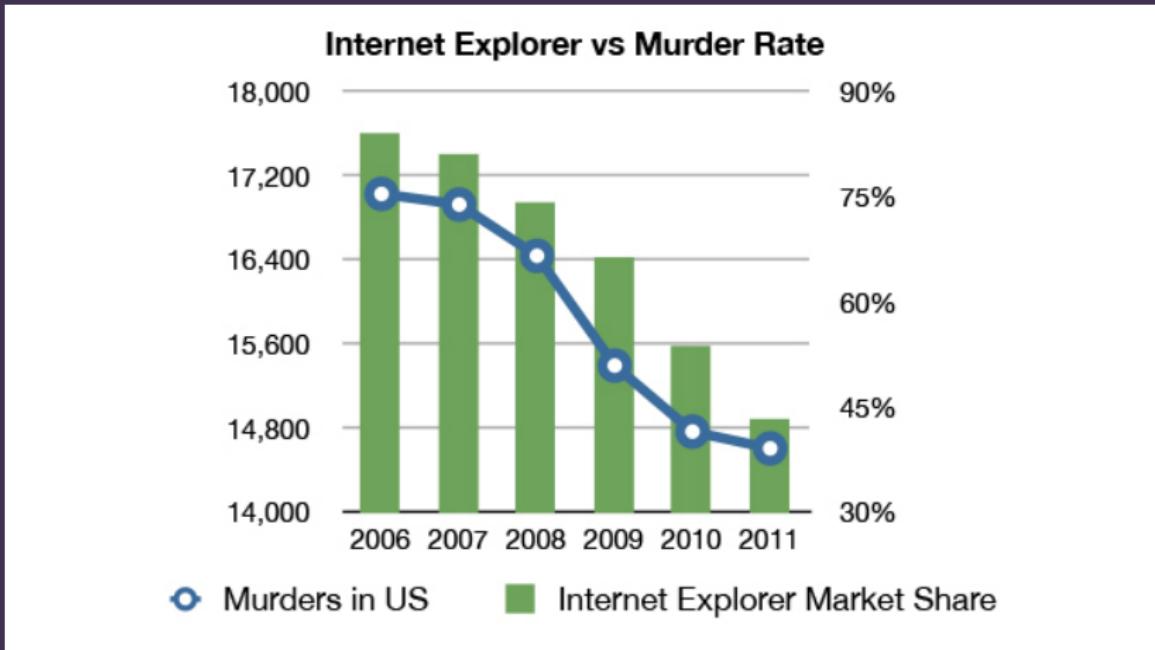


# Correlation does not imply causation

# Correlation does not imply causation



# Correlation does not imply causation



Data-driven decision making





<https://youtu.be/HQwL6zh7AgA>  
[http://media.steampowered.com/apps/  
steamdevdays/slides/data.pdf](http://media.steampowered.com/apps/steamdevdays/slides/data.pdf)

# Decision making at Valve

# Decision making at Valve

- ▶ Valve favour **data-driven decision making**

# Decision making at Valve

- ▶ Valve favour **data-driven decision making**
  - ▶ Ask **explicit** questions

# Decision making at Valve

- ▶ Valve favour **data-driven decision making**
  - ▶ Ask **explicit** questions
  - ▶ Look at the **data**

# Decision making at Valve

- ▶ Valve favour **data-driven decision making**
  - ▶ Ask **explicit** questions
  - ▶ Look at the **data**
  - ▶ Use the data to develop a **theory**

# Decision making at Valve

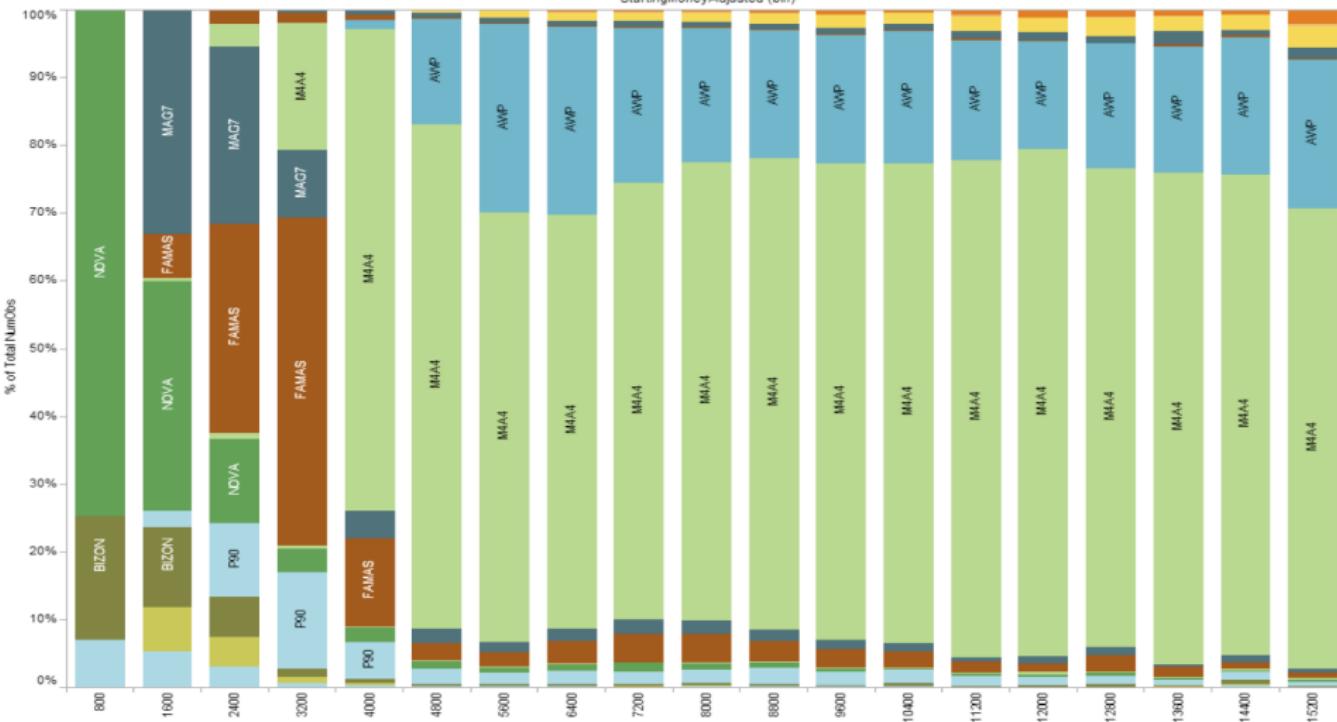
- ▶ Valve favour **data-driven decision making**
  - ▶ Ask **explicit** questions
  - ▶ Look at the **data**
  - ▶ Use the data to develop a **theory**
  - ▶ Define **measurable outcomes**

# Decision making at Valve

- ▶ Valve favour **data-driven decision making**
  - ▶ Ask **explicit** questions
  - ▶ Look at the **data**
  - ▶ Use the data to develop a **theory**
  - ▶ Define **measurable outcomes**
  - ▶ **Iterate**

### Weapon Purchases by Budget

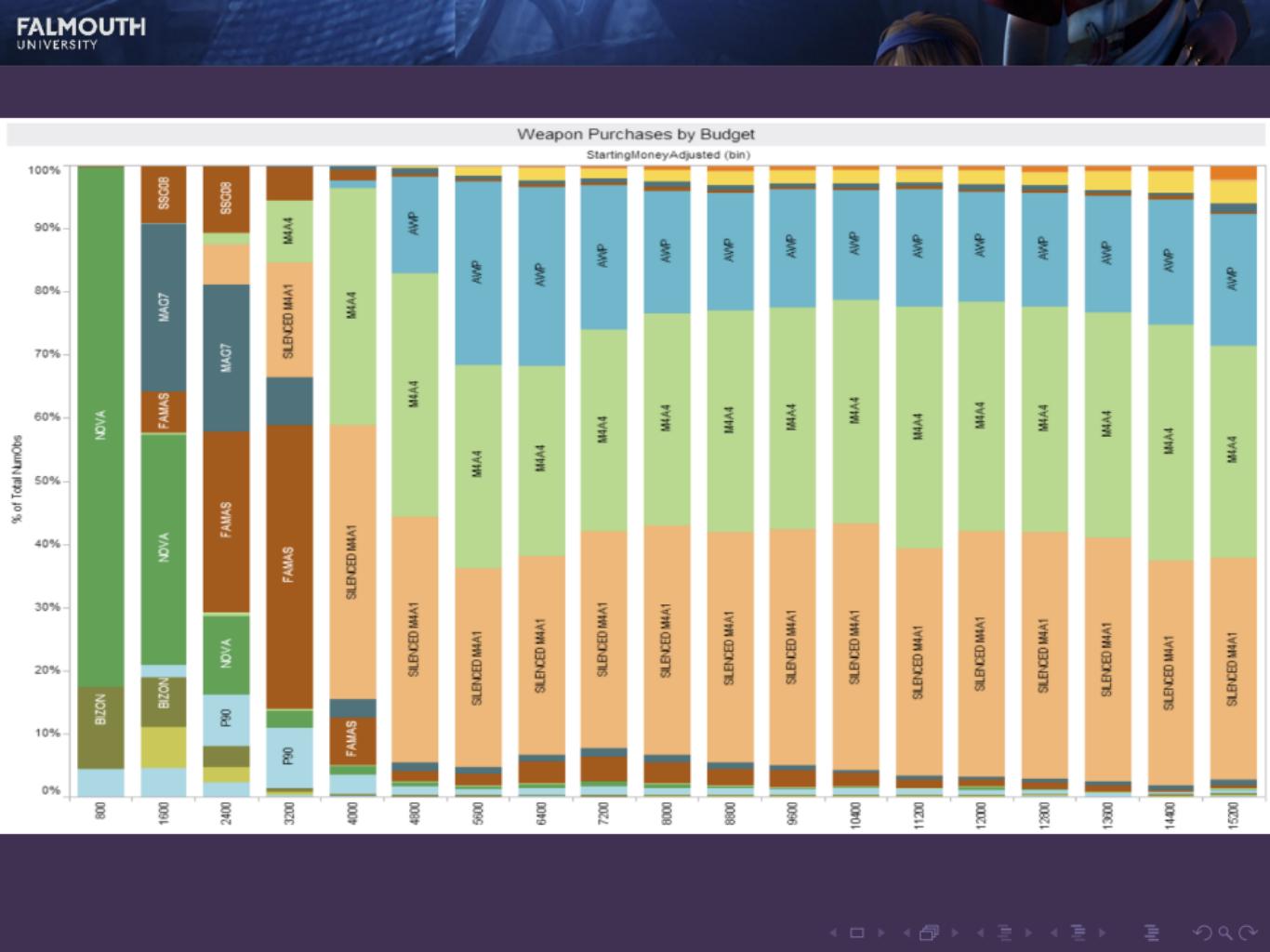
StartingMoneyAdjusted (bin)



# Weapon Balance

- Explicit: M4A4 usage is high; few choices in late-game
- Data-driven: Purchase rates
- Theory-driven: Greater tactical choice → Player retention
- Measurements: Purchase rates, playtime, efficacy
- Iterative: Inform future design choices

Hypothesis: Creating a balanced alternative weapon will increase player choice and playtime



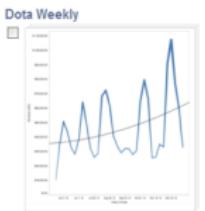
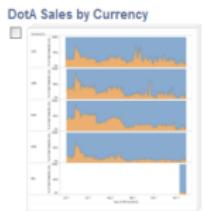
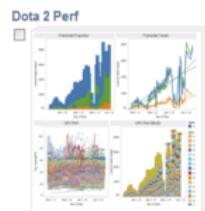
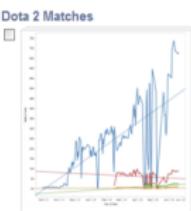
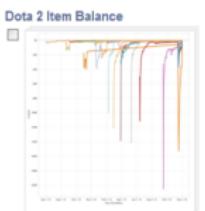
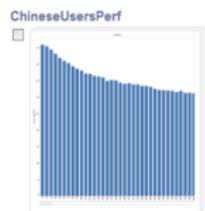
# Results

- ~ 50/50 split between new and old favorites
- Increase in playtime
  - Conflated with other updates
  - Difficult to isolate
- Open question as to whether or not increased weapon variability increases player retention

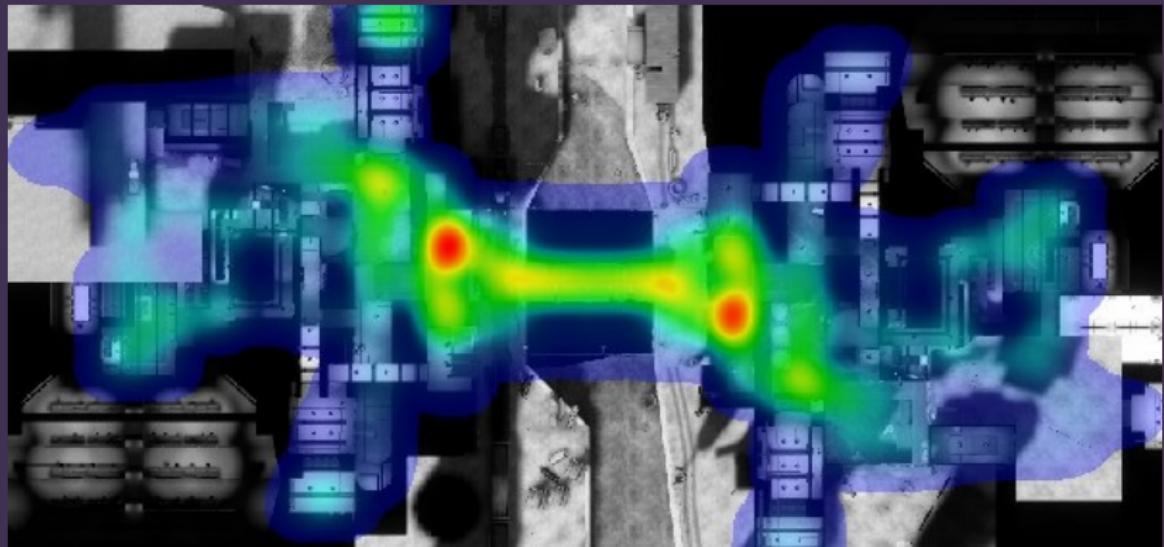
# Data visualisation



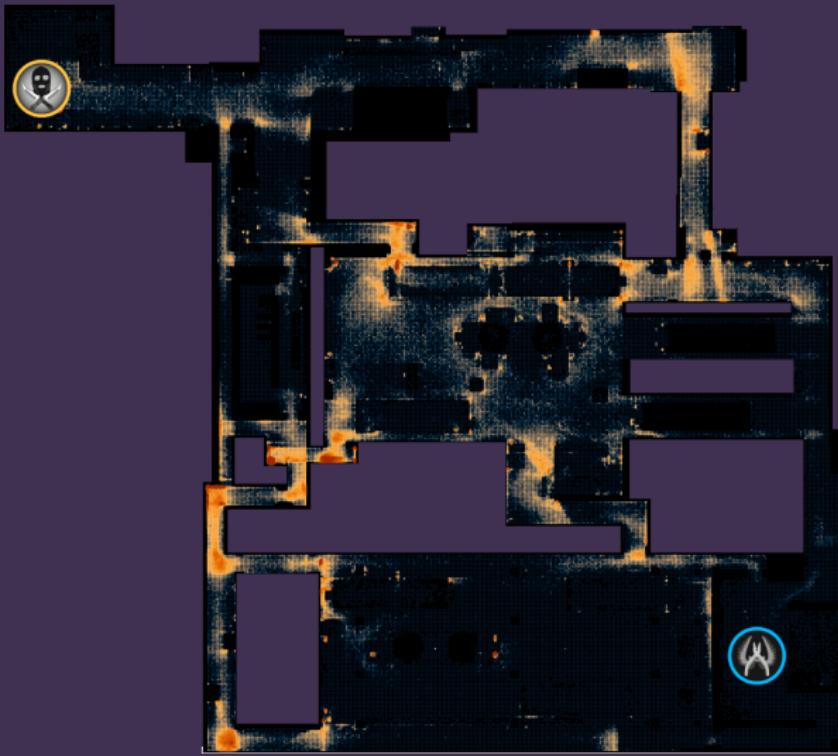
## ValveStats



# Player deaths in Team Fortress 2



# Weapon fire locations in CS:GO



# Data visualisation

# Data visualisation

- Humans are good at seeing **patterns**

# Data visualisation

- ▶ Humans are good at seeing **patterns**
- ▶ Good data visualisation can help to spot patterns

# Data visualisation

- ▶ Humans are good at seeing **patterns**
- ▶ Good data visualisation can help to spot patterns
- ▶ ... However this should be followed up by proper statistical analysis!

# Psychology





<http://gdcvault.com/play/1017940/>

The-Science-Behind-Shaping-Player

<https://backchannel.com/>

inside-the-largest-virtual-psychology-lab-in-the-world  
.e63is9hkl

# Priming

# Priming

- ▶ How do you pronounce this word: S - H - O - P?

# Priming

- ▶ How do you pronounce this word: S - H - O - P?
- ▶ What do you do at a green traffic light?

# A/B testing

# A/B testing

- ▶ Push **different versions** of your game to different groups of players

# A/B testing

- ▶ Push **different versions** of your game to different groups of players
- ▶ Measure over different groups to compare effects of changes

# A/B testing

- ▶ Push **different versions** of your game to different groups of players
- ▶ Measure over different groups to compare effects of changes
- ▶ Not restricted to two versions

# Priming in League of Legends

# Priming in League of Legends

- ▶ **Hint text** on loading screen and/or in-game

# Priming in League of Legends

- ▶ **Hint text** on loading screen and/or in-game
- ▶ “X% of players punished by the Tribunal improve their behaviour and are never punished again”

# Priming in League of Legends

- ▶ **Hint text** on loading screen and/or in-game
- ▶ “X% of players punished by the Tribunal improve their behaviour and are never punished again”
- ▶ Led to a **6% decrease** in verbal abuse and offensive language

# Priming in League of Legends

# Priming in League of Legends

- ▶ “Teammates perform worse if you harass them after a mistake”

# Priming in League of Legends

- ▶ “Teammates perform worse if you harass them after a mistake”
- ▶ → **No significant** impact

# Priming in League of Legends

- ▶ “Teammates perform worse if you harass them after a mistake”
- ▶ → **No significant** impact
- ▶ “Teammates perform worse if you harass them after a mistake”

# Priming in League of Legends

- ▶ "Teammates perform worse if you harass them after a mistake"
- ▶ → **No significant impact**
- ▶ "Teammates perform worse if you harass them after a mistake"
- ▶ → **11% decrease** in offensive language

# Priming in League of Legends

# Priming in League of Legends

- ▶ “Players who cooperate with their teammates win  $X\%$  more games”

# Priming in League of Legends

- ▶ “Players who cooperate with their teammates win  $X\%$  more games”
- ▶ → **No significant** impact

# Priming in League of Legends

- ▶ “Players who cooperate with their teammates win X% more games”
- ▶ → **No significant** impact
- ▶ “Players who cooperate with their teammates win X% more games”

# Priming in League of Legends

- ▶ “Players who cooperate with their teammates win X% more games”
- ▶ → **No significant** impact
- ▶ “Players who cooperate with their teammates win X% more games”
- ▶ → **6% decrease** in offensive language

# Priming in League of Legends

# Priming in League of Legends

- ▶ “Who will be the most sportsmanlike player in the game?”

# Priming in League of Legends

- ▶ “Who will be the most sportsmanlike player in the game?”
- ▶ → **15% increase** in offensive language

# Games as experiments

# Games as experiments

- ▶ These experiments would be difficult to do in the lab

# Games as experiments

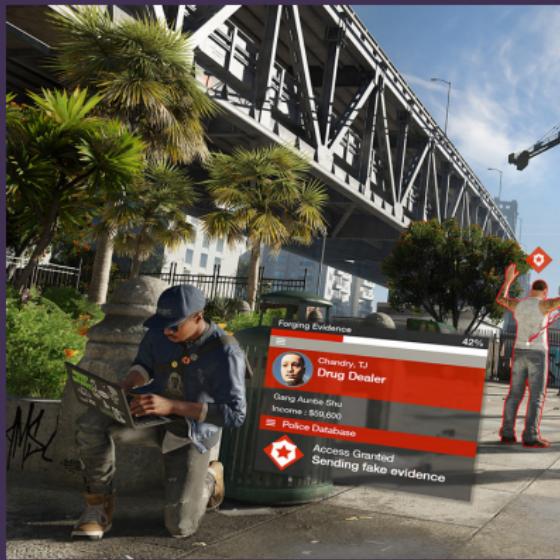
- ▶ These experiments would be difficult to do in the lab
- ▶ ... and impossible at this scale (**millions** of participants)

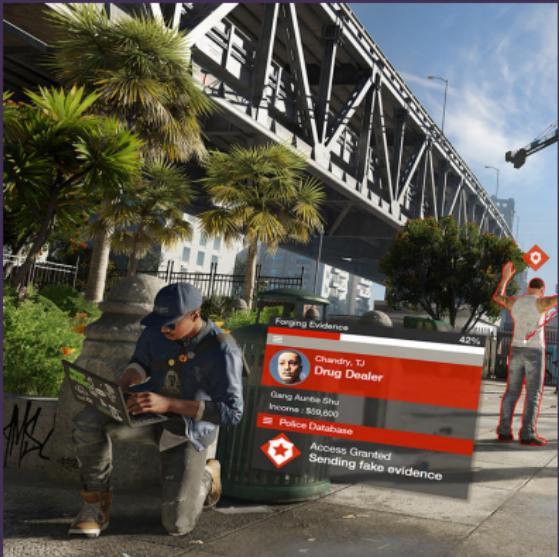
# Games as experiments

- ▶ These experiments would be difficult to do in the lab
- ▶ ... and impossible at this scale (**millions** of participants)
- ▶ Results have a **measurable positive impact** on player experience

# Ethics of analytics







through our carriers or network service providers.

b. Collection of personal data: In order to provide You with a better game experience, adapted services and Product support, UBISOFT may collect and store data about You in relation to Your use of the Product, Your connection information and/or Your Compatible Mobile Terminal. Certain data is recorded, archived, analysed and used to create user statistics. Your privacy is very important to UBISOFT and UBISOFT will not reveal Your personal data to third parties except when expressly authorised by You to do so or in special circumstances. UBISOFT may be under a duty to disclose or share Your personal data in order to comply with a legal obligation, or in order to protect UBISOFT's rights and those of other users and third parties. This includes exchanging information with other companies and organisations for the purposes of fraud protection and credit risk reduction. In addition, UBISOFT reserves the right to collect, store and use anonymous data about You. For further information concerning UBISOFT's use of Your personal data, please refer to the Privacy Policy.

c. Analytics Tools and Ad Serving Technology. UBISOFT uses third party analytics tools to collect information concerning Your and other users' gaming habits and use of the Product. The information collected may contain the following, without limitation: mobile device unique identity or other device identifiers and settings, carrier, operating system, localization information, date and time spent on the Product, game scores, game metrics and statistics, feature usage, advertising conversion rates, monetization rate, purchase history and other similar information. UBISOFT uses third party ad serving technologies that may collect information as a result of ad serving in the Product and that may temporarily display advertisements in the Product. The information collected and used for the targeted advertising may contain the following, without limitation: age and gender, number views of an advertisement, mobile device unique identity or other device identifiers and settings, information about Your use of the Product, Compatible Mobile Terminal, and other Internet sites (including third party web pages and mobile Internet sites) viewed by You (as well as date and time viewed), advertisement(s) served, the advertisement(s)' in game location and length, and Your response to the advertisement(s) (if any). The analytics tools and ad serving technologies may use server log files, web beacons, cookies, tracking pixels and other technologies to collect said information and may combine the information collected on other UBISOFT Products and services with information collected from other third party websites and mobile products and services and with demographic, advertisement, market and other analytics surveys. Standing alone, this information is not personal data; however, if UBISOFT combines any of this information with personal data, UBISOFT will treat this information as personal data pursuant to our Privacy Policy. You will find a non-exhaustive list of our partners whose analytic tools and ad serving technologies are used in our Product(s) at <https://legal.ubi.com/partners>. UBISOFT makes its best efforts to provide You with the links to our partners' own privacy policies according to which their tools and technologies are used and when available the link to allow You to opt-out from their services.

#### 4. CONSENT TO MONITOR.

When You are using the Product, the Product may monitor Your hardware random access memory (RAM) for unauthorized third party programs prohibited by Section 1 that interact with the Product. In the

# Ethical considerations

# Ethical considerations

- Are you breaching your players' **privacy**?

# Ethical considerations

- ▶ Are you breaching your players' **privacy**?
- ▶ Is it right to **experiment** on your players?

# Ethical considerations

- ▶ Are you breaching your players' **privacy**?
- ▶ Is it right to **experiment** on your players?
- ▶ Is it right to **manipulate** your players' behaviour?

# Ethical considerations

- ▶ Are you breaching your players' **privacy**?
- ▶ Is it right to **experiment** on your players?
- ▶ Is it right to **manipulate** your players' behaviour?
- ▶ Are you deliberately adding **addictive** qualities to your game?

# Ethical considerations

- ▶ Are you breaching your players' **privacy**?
- ▶ Is it right to **experiment** on your players?
- ▶ Is it right to **manipulate** your players' behaviour?
- ▶ Are you deliberately adding **addictive** qualities to your game?
- ▶ Is the above justified if it improves the player **experience**?

# Ethical considerations

- ▶ Are you breaching your players' **privacy**?
- ▶ Is it right to **experiment** on your players?
- ▶ Is it right to **manipulate** your players' behaviour?
- ▶ Are you deliberately adding **addictive** qualities to your game?
- ▶ Is the above justified if it improves the player **experience**?
- ▶ What if it improves your **profits** instead / as well?

# Legal considerations

# Legal considerations

- ▶ **NB: this slide is for education only and does NOT constitute legal advice!**

# Legal considerations

- ▶ NB: this slide is for education only and does NOT constitute legal advice!
- ▶ Legislation such as the **General Data Protection Regulation (GDPR)** and the **Data Protection Act**

# Legal considerations

- ▶ **NB: this slide is for education only and does NOT constitute legal advice!**
- ▶ Legislation such as the **General Data Protection Regulation (GDPR)** and the **Data Protection Act**
- ▶ Covers **personal data**: any data that can be used to identify a living individual

# Legal considerations

- ▶ **NB: this slide is for education only and does NOT constitute legal advice!**
- ▶ Legislation such as the **General Data Protection Regulation (GDPR)** and the **Data Protection Act**
- ▶ Covers **personal data**: any data that can be used to identify a living individual
  - ▶ Name, phone number, email address, IP address, social media ID, ...

# Legal considerations

- ▶ NB: this slide is for education only and does NOT constitute legal advice!
- ▶ Legislation such as the **General Data Protection Regulation (GDPR)** and the **Data Protection Act**
- ▶ Covers **personal data**: any data that can be used to identify a living individual
  - ▶ Name, phone number, email address, IP address, social media ID, ...
- ▶ Covers the **processing** (including storage) of personal data

# Legal considerations

- ▶ NB: this slide is for education only and does NOT constitute legal advice!
- ▶ Legislation such as the **General Data Protection Regulation (GDPR)** and the **Data Protection Act**
- ▶ Covers **personal data**: any data that can be used to identify a living individual
  - ▶ Name, phone number, email address, IP address, social media ID, ...
- ▶ Covers the **processing** (including storage) of personal data
- ▶ The data processor has certain **responsibilities**

# Legal considerations

- ▶ NB: this slide is for education only and does NOT constitute legal advice!
- ▶ Legislation such as the **General Data Protection Regulation (GDPR)** and the **Data Protection Act**
- ▶ Covers **personal data**: any data that can be used to identify a living individual
  - ▶ Name, phone number, email address, IP address, social media ID, ...
- ▶ Covers the **processing** (including storage) of personal data
- ▶ The data processor has certain **responsibilities**
- ▶ The data subject has certain **rights**

# Legal considerations

- ▶ NB: this slide is for education only and does NOT constitute legal advice!
- ▶ Legislation such as the **General Data Protection Regulation (GDPR)** and the **Data Protection Act**
- ▶ Covers **personal data**: any data that can be used to identify a living individual
  - ▶ Name, phone number, email address, IP address, social media ID, ...
- ▶ Covers the **processing** (including storage) of personal data
- ▶ The data processor has certain **responsibilities**
- ▶ The data subject has certain **rights**
- ▶ Not complying can be a **civil and/or criminal offence**

# Conclusion

# Conclusion

- ▶ Analytics can be seen as **large-scale playtesting**

# Conclusion

- ▶ Analytics can be seen as **large-scale playtesting**
- ▶ Allows a **scientific** approach to game design decisions

# Conclusion

- ▶ Analytics can be seen as **large-scale playtesting**
- ▶ Allows a **scientific** approach to game design decisions
- ▶ Allows a **scientific** approach to business decisions

