

People and Spam

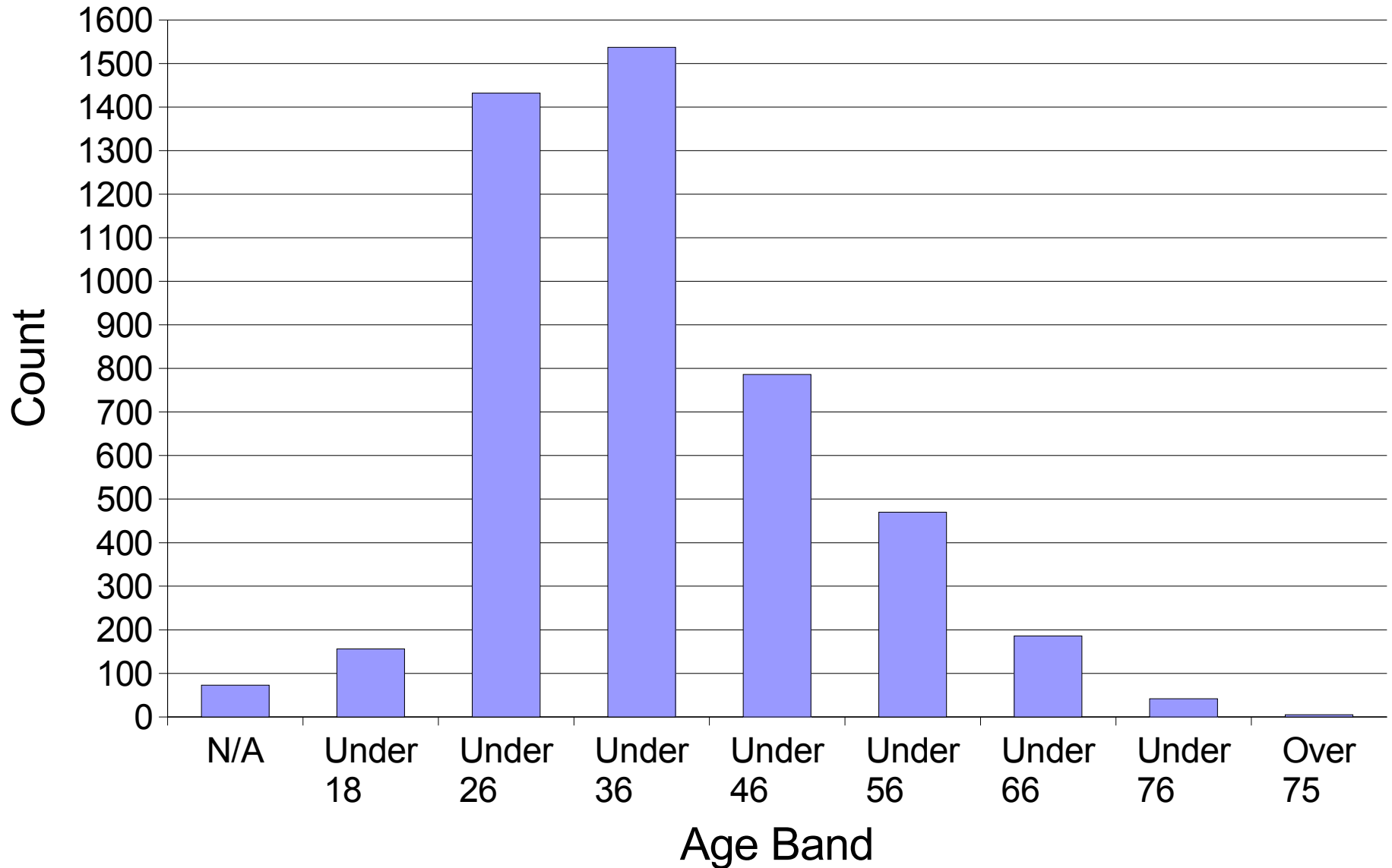
John Graham-Cumming
The POPFile Project & Electric Cloud, Inc.
January 21, 2005



Largest End-user Survey

- Get real non-vendor data
- 4,691 participants
 - 94.5% men, 4% women, 1.5% other
 - Experienced email users: 42% > 10 years, 50% 6 to 10 years.
 - Experienced users: 37% experienced computer users, 36% programmers, 23% sysadmins.
 - 2 respondents identified themselves as spammers
 - Both claim to use and antispam tool
 - 80% of respondents between 18 and 45.

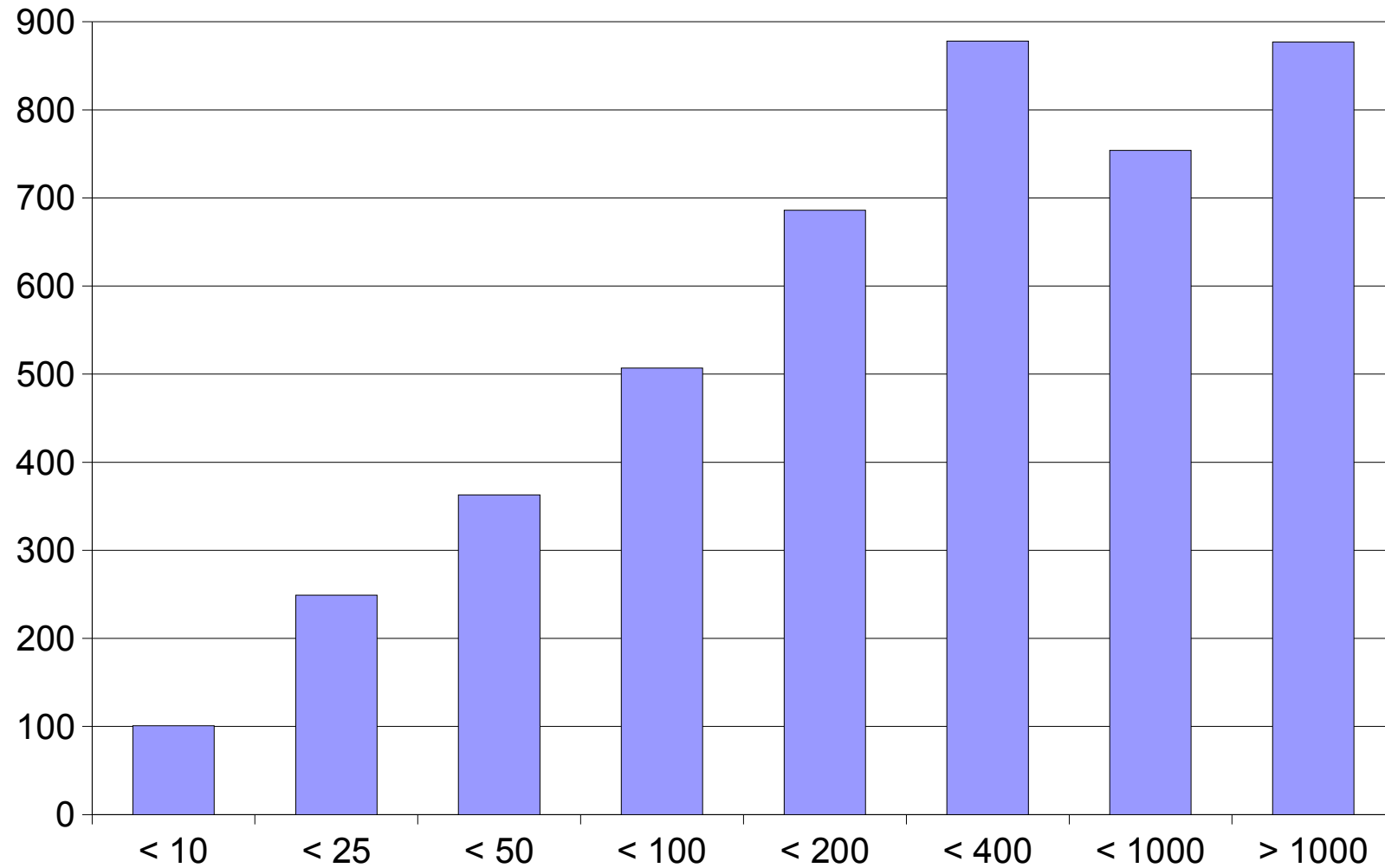
Age Groups



General

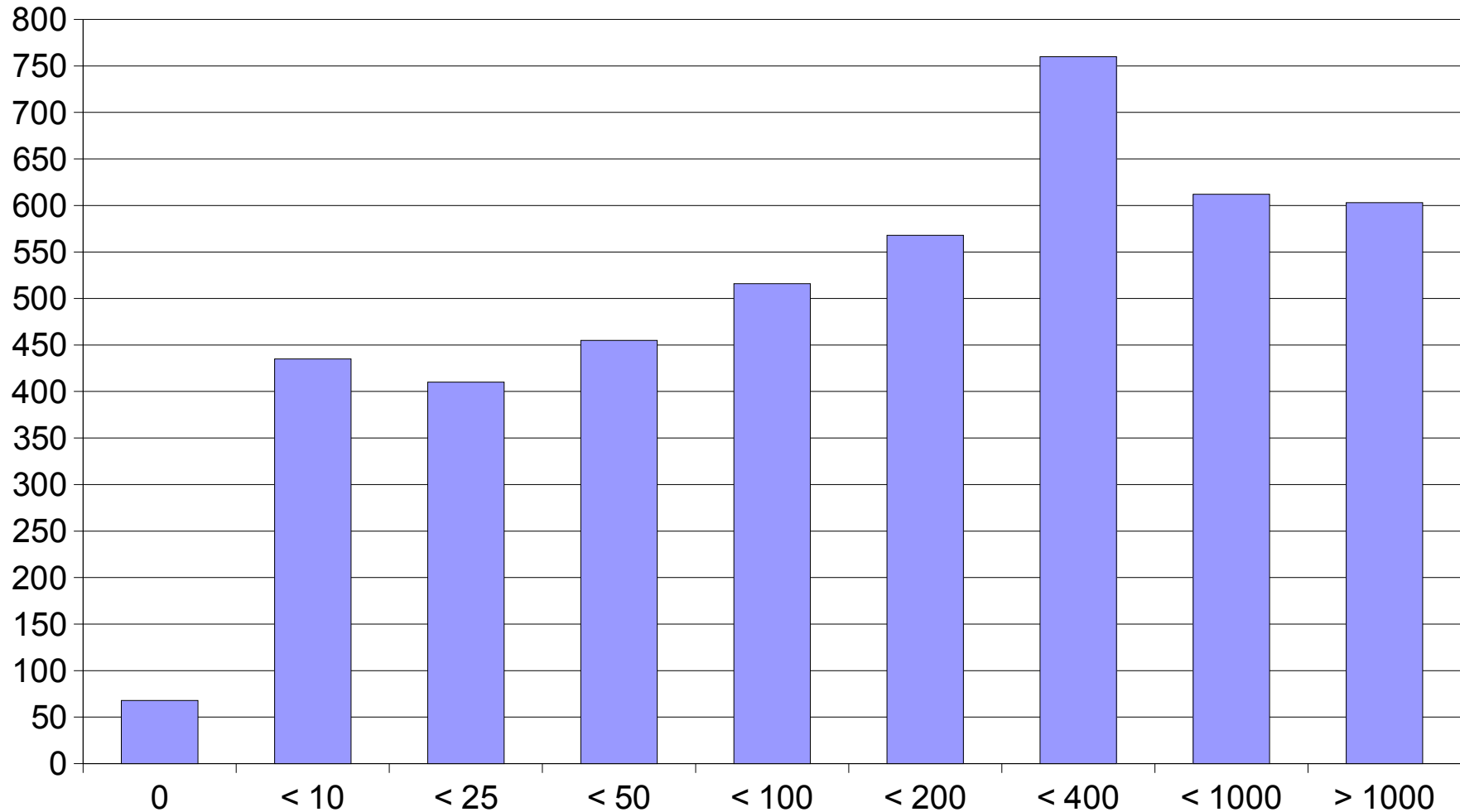
- 76% believe that the spam problem will never go away (women more pessimistic than men)
- 9% believe spam filtering makes the spam problem worse
- 85% are using an anti-spam tool (same for men and women)
- All the source data from this survey is available at <http://www.jgc.org/>

Mail Messages Per Week



- Average user get 413 mails per week

Spam Messages Per Week



- 98.5% of users receive spam
- Average user receives 318 spams per week
- Indicates 77% of mail is spam

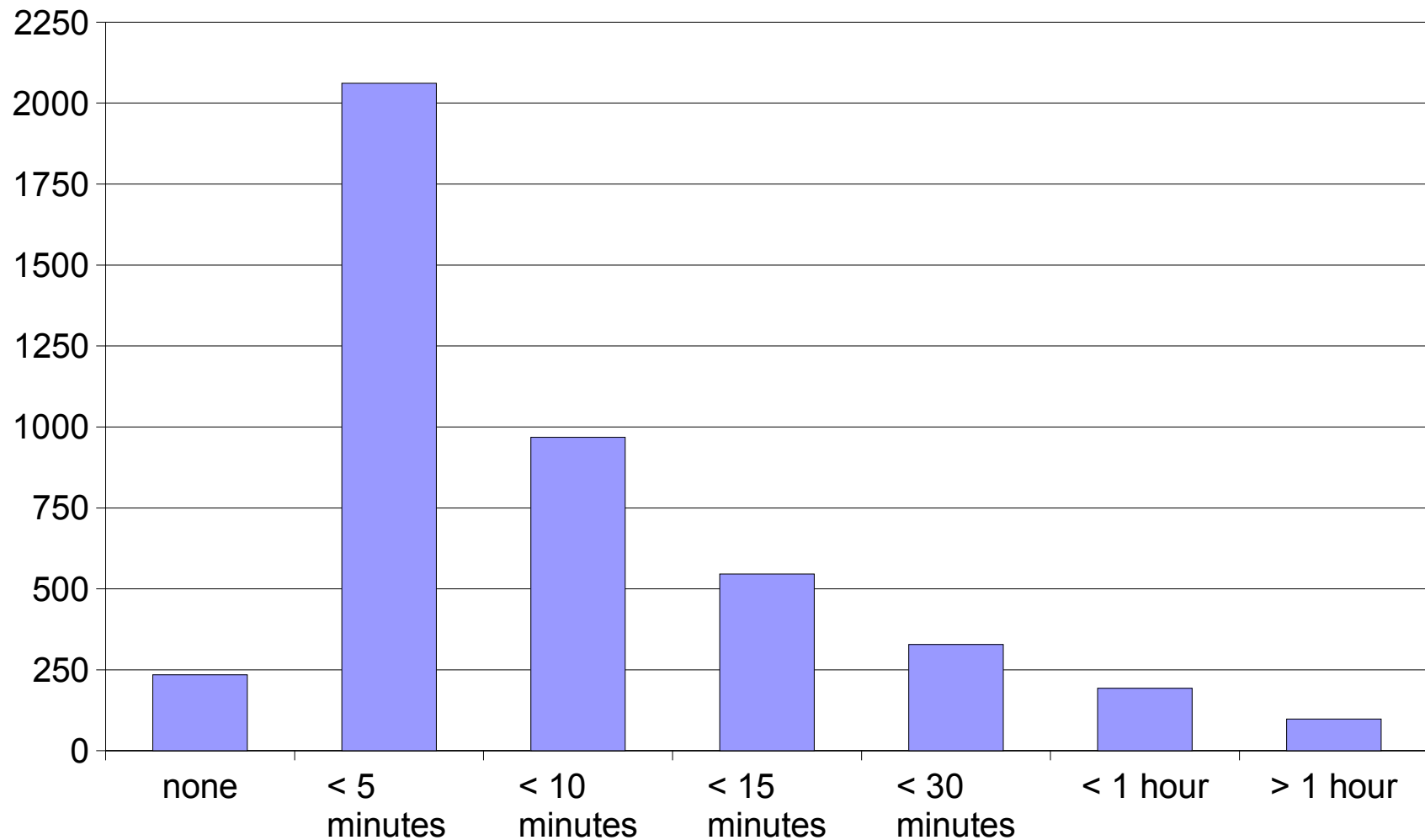
Who's buying?

- Two questions asked about buying from spam and buying from regular, physical junk mail:
 - 1% of respondents say that they have bought from spam
 - 16% have bought from physical junk mail
 - No difference between men and women
- So, even highly experienced computer and email users buy from spam
- 15% of users are not using any antispam software

Of people who buy from spam...

- Much more likely to buy from physical junk mail (70% vs 16%)
- Consider spam to be less of a problem than others
 - 10% don't care about spam vs 1.7%
 - Less likely to be annoyed by spam (44% vs 50%)
 - Half as likely to feel violent towards spammers
- More likely to have signed up for pornography (27% vs 10%)
- More likely to pay for an anti-spam solution
- 25% are between 46 and 55; same level of experience

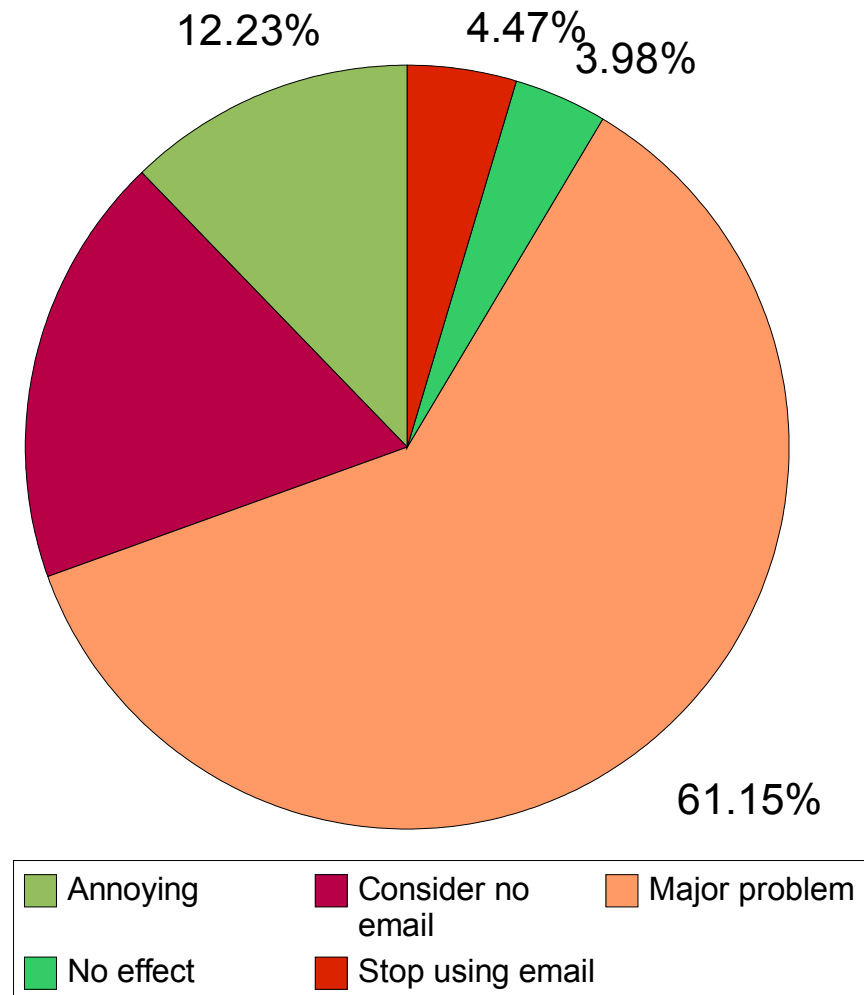
Time Dealing with Spam



- 1/3 of users spend 5 to 30 minutes a day dealing with spam
- 50% spend under 5 minutes
- On average spend 9 mins/day dealing with spam (men and women)

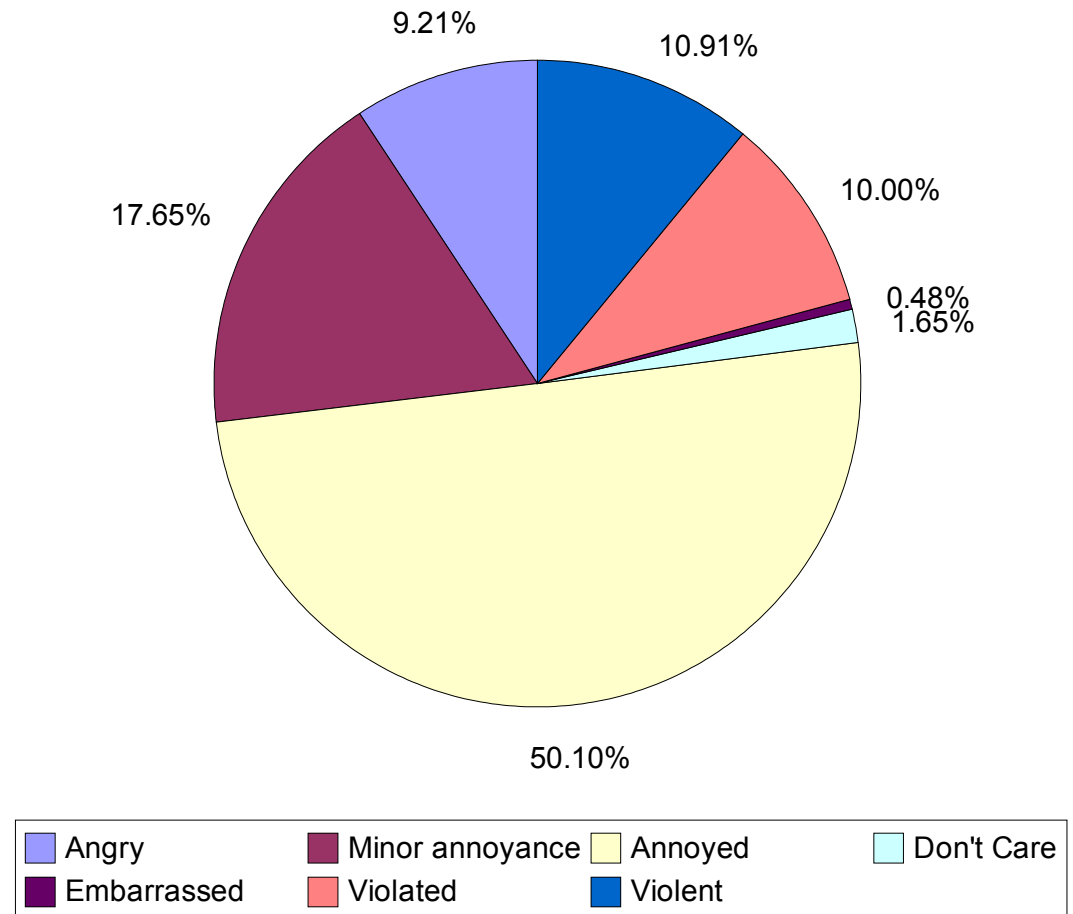
The effect of spam

- Without a spam filter...
 - 23% would stop or considering stopping all use of email (equal for men and women)
 - 61% would have major problems using email
 - 4% wouldn't have a problem

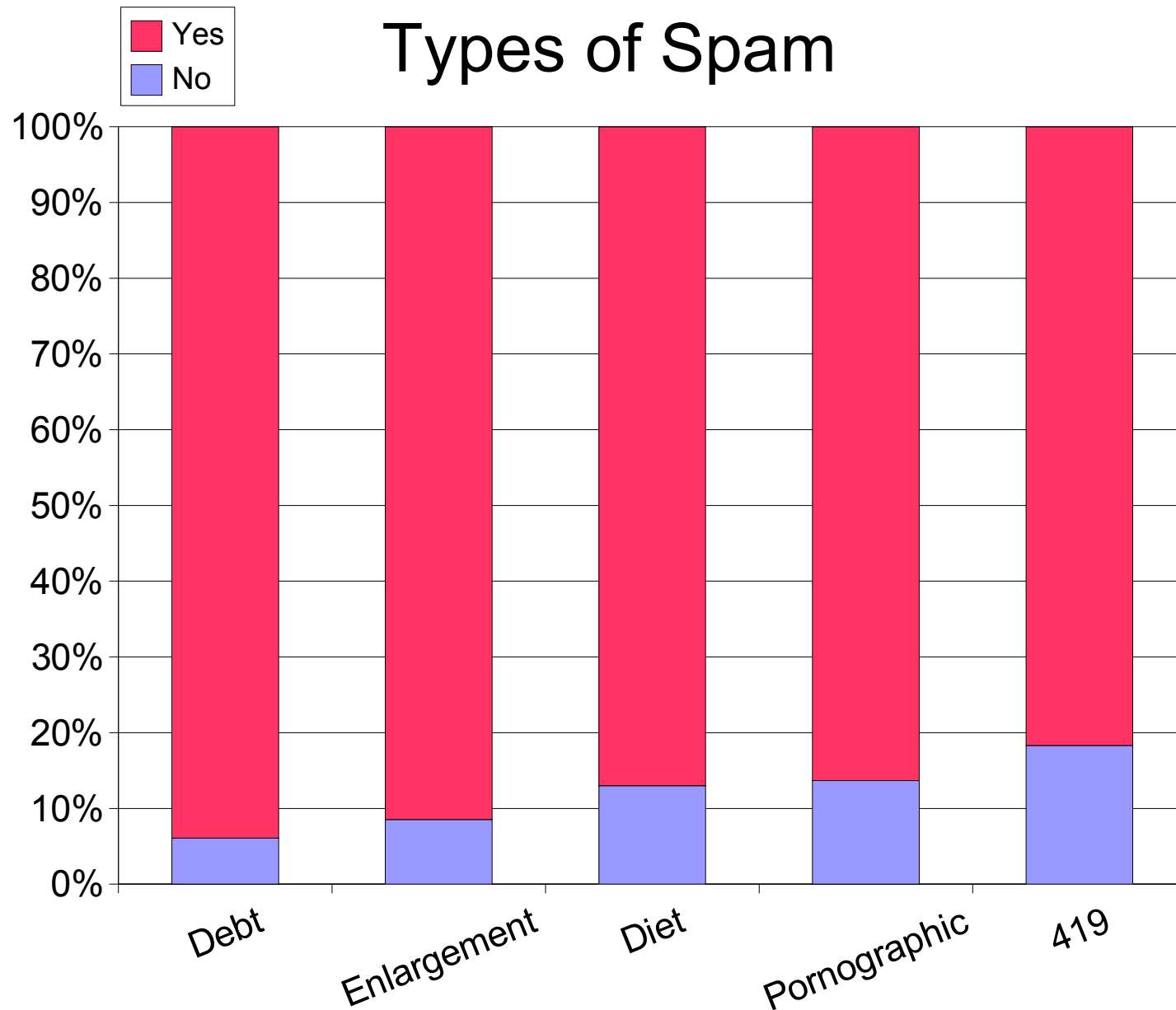


The effect of spam

- For most people spam is a daily annoyance
- 11% would like to physically harm a spammer (women half as violent as men)
- 10% feel their private space has been violated (women twice 50% more likely to feel violated)
- Only 1.7% aren't bothered at all; spam affects 98.3%



Types of Spam



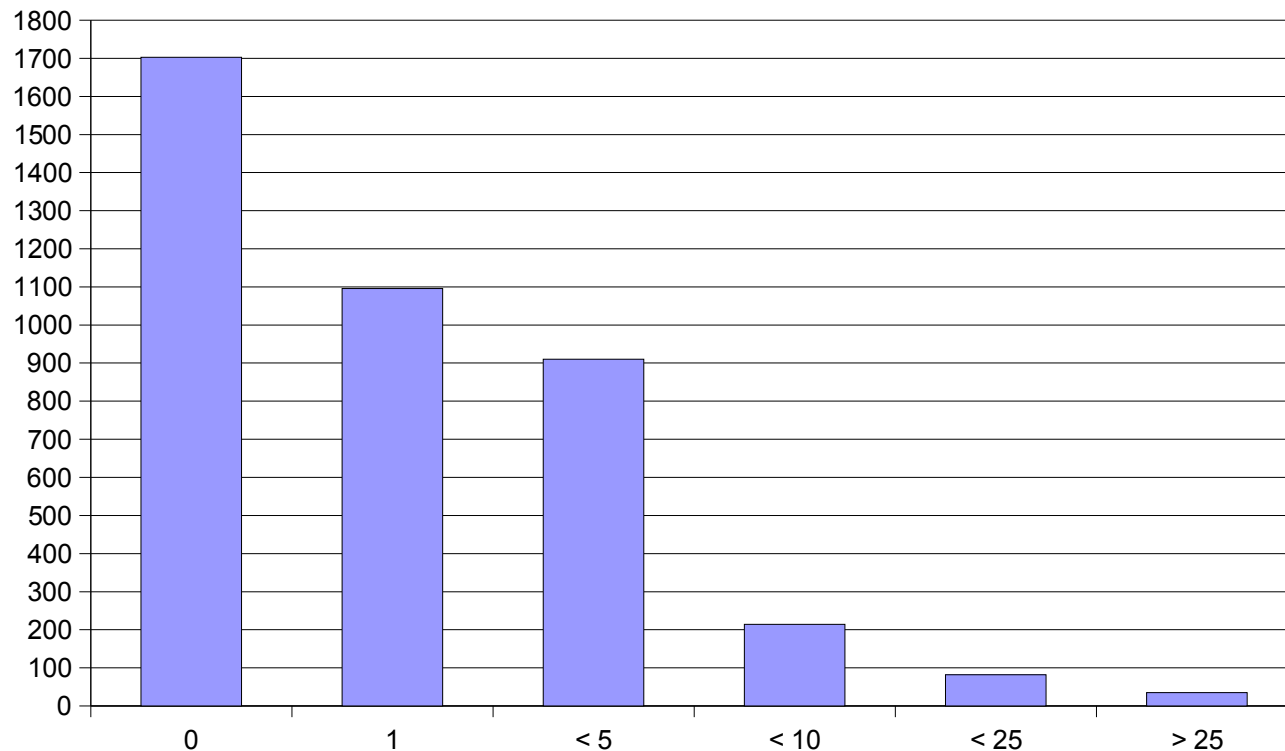
- 10% have given their email address to a pornographic site (men twice as likely as women); yet, 86% receive pornographic spam

Children and Older People

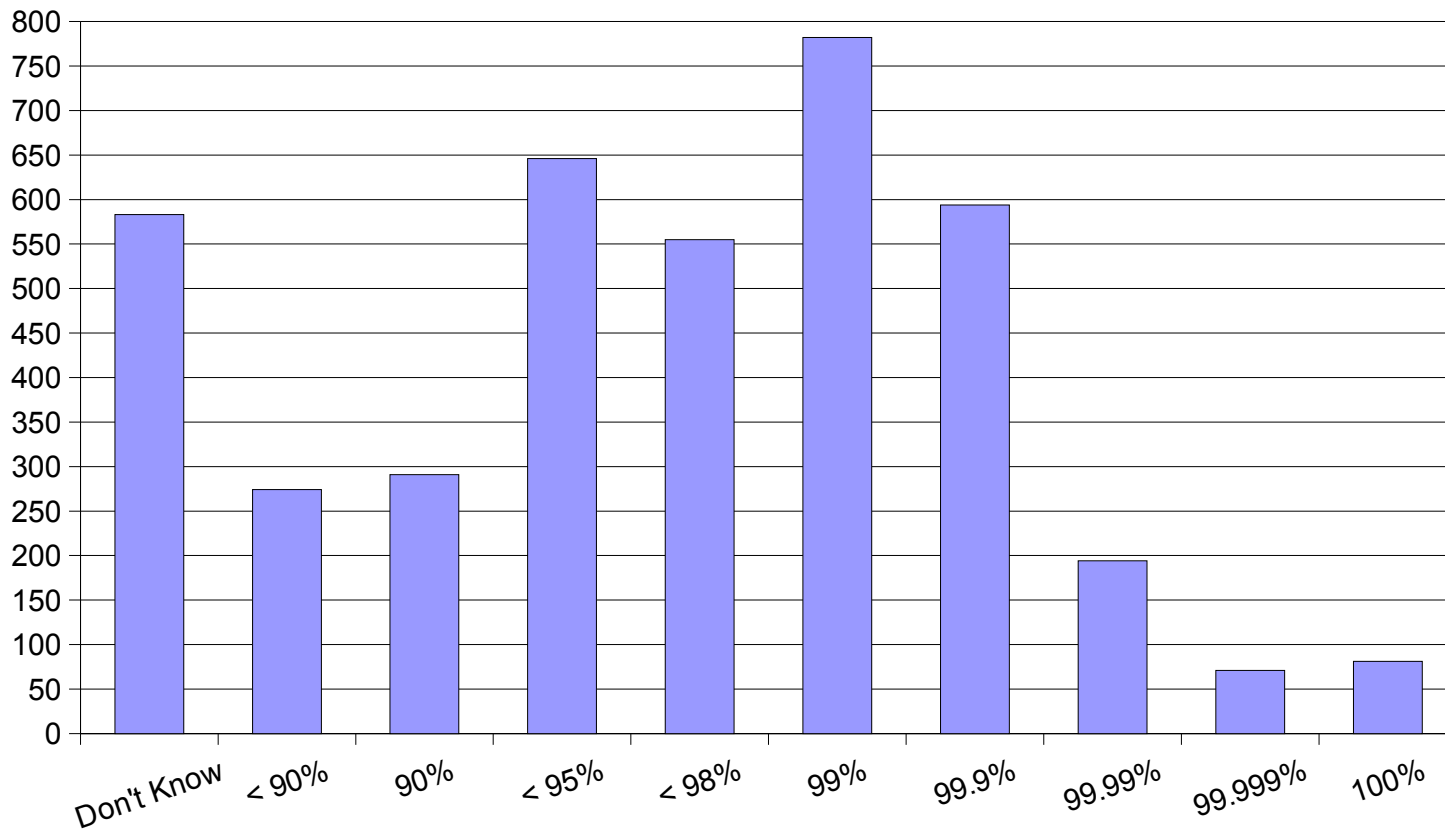
- 15% of respondents have children
 - 74% of children receive spam
 - 60% of children receive pornographic spam
- 5% of respondents are over 55
 - Take 20% longer dealing with spam (11m vs 9m)
 - An average amount of mail (422 per week)
 - 2% have bought from spam (double the average)
- 3% of respondents under 18
 - Less likely to use anti-spam tool (75% have tool)
 - 11% say have signed up with pornographic site (bad data?)

Anti-spam Tools

- 42% say they have lost mail because of anti-spam
- False positives per week (42% get none):



Anti-spam Tool Accuracy



- 42% getting 99% accuracy or better

Testing Human Accuracy

- Survey also tested human spam filtering accuracy
- People had to manually sort 260 messages
- Messages were 80% spam
- Only the From and Subject were visible
- Messages were from my real email
- People had to pretend to be me

Human Accuracy

- Overall accuracy was 99.46%
 - 4x as likely to think a good message was spam than vv
 - Similar for men and women, and all age groups
- Test showed messages in groups of 1, 5, 10, 20 and 50
 - No significant differences in accuracy between them
 - Get 100% accuracy if one message presented at a time
- Test showed messages unsorted, sorted by From and sorted by Subject
 - No significant differences in accuracy between them
 - Slightly better if messages were unsorted

Caveats

- Didn't attempt to validate people's answers
 - e.g. no checking of ages or sexes
- No attempt to provide statistically meaningful data for the general population
- Data is skewed heavily to men, between 18 and 45 who are experienced computer users
- I am not a professional pollster

Hot Free Stuff!

- The Spammers' Compendium
 - <http://www.jgc.org/tsc/>
- jgc's spam and anti-spam newsletter
 - Twice per month
 - <http://www.jgc.org/>
- POPFile
 - Sorts more than spam
 - <http://getpopfile.org/>