



Double Blind Coffee Taste Test

Results

MC

Summary and Statistics

The Study

On the 1st of September 2018, five intrepid scientists (Emma, Evan, Faye, Maciek and Spencer) brewed and sampled 11 different coffees, to determine which was their favourite.

By using a pseudo-random number generator and two independent individuals assigning codes to each coffee, they implemented a double-blind design so no-one could know which coffee they were drinking. The coffees were drunk in a random order, but the same random order for all participants. Each coffee was brewed in one of two only slightly different stove top percolators.

Four of the 11 coffees came pre-ground, while the remainder were ground fresh on the day. Five of the coffees were able to be classified a roast type, among “Light”, “Medium”, “Medium-Dark” and “Dark”. The remaining coffees were either decaf (two of them) or too pretentious to have roast published. Each coffee was also assigned a price based on its current online store price in CA\$.

Each coffee was rated on:

Aroma How the coffee smelled

Taste Drinkability and “general goodness”

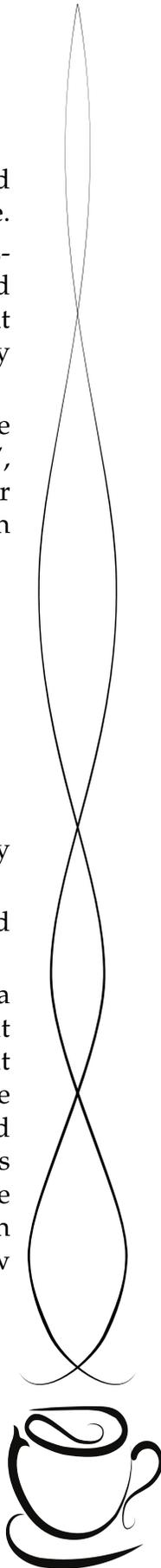
Aftertaste However the participants wanted to interpret this

Each rating was on the scale: Very Poor (-2), Poor (-1), Medium (0), Good (1), Very Good (2).

Each participant also tried to guess which coffee they were drinking, and then typed out a pretentious description of the coffee.

The results were statistically analysed as linear models fit by MLE, with each rater as a random grouping effect on the intercept alone. In English, that means that I assumed that each person had a different baseline rating—some people just like coffee more—but that we can still pool information from everyone’s ratings to understand something about the coffees themselves, like how much better a coffee gets per dollar in price. The way I did that was, basically, by lining up all ratings against some variable (for example: Price Vs Taste), and then drawing a straight line through that cloud of points, trying to get the line as close as possible to every single point. Just like that, except I drew the line with math instead of on paper. How steep that line is gives us some insight into, for example, how the price of a coffee affects its Taste rating.

This scientific adventure yielded the following important discoveries.



No evidence of boring spurious confound effects

There was no evidence that either the pot the coffee was brewed in ($p = 0.5$) nor the order it was drunk in ($p = 0.7$) influenced rater's opinions of the coffee. Given our tiny sample size, these potential confounds were discarded (rather than being controlled for in the substantive models) to avoid sapping our already puny statistical power.

Price Effect

Overall, across all raters and all coffees:

Aroma Increased by 0.06 rating points per dollar; or, in other words, it cost about \$16 for a 1 rating point increase.

Taste Also increased by 0.06 rating/dollar.

Aftertaste Increased by 0.07 rating/dollar; so it only cost \$13 for a 1 point aftertaste improvement.

Pre-ground Effect

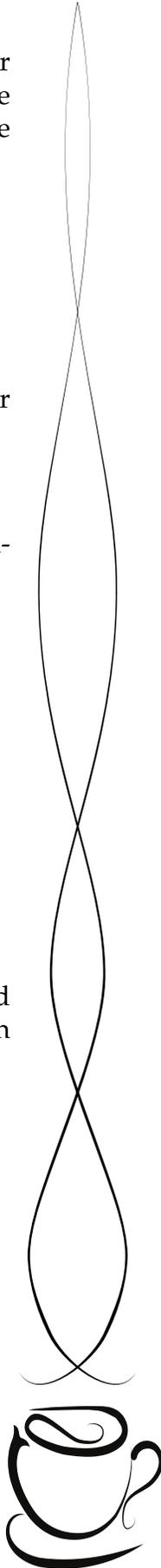
Overall, across all raters and all coffees:

Aroma Fresh ground coffees were rated 0.5 points higher.

Taste Fresh ground coffees were rated 0.9 points higher.

Aftertaste Fresh ground coffees were rated 0.7 points higher.

These results control for the price effect. That means, in English, that even over and above how much better you'd expect a coffee to be because it is more expensive, fresh ground coffees are almost a full rating point better still.



Guessability

Of the 55 guesses taken by the raters, only 7 correctly guessed the coffees they were drinking. This is not statistically different from random dumb luck—just guessing at random would produce at least this many correct answers 12% of the time.

On the other hand, 24 of the guesses at least were correct about the brand of coffee (i.e., Kicking Horse, Pilot, etc.). There’s only about a one in ten million chance of this many accurate guesses occurring by luck alone, so it seems like the raters at least had some inkling of the brand of coffee they were drinking.

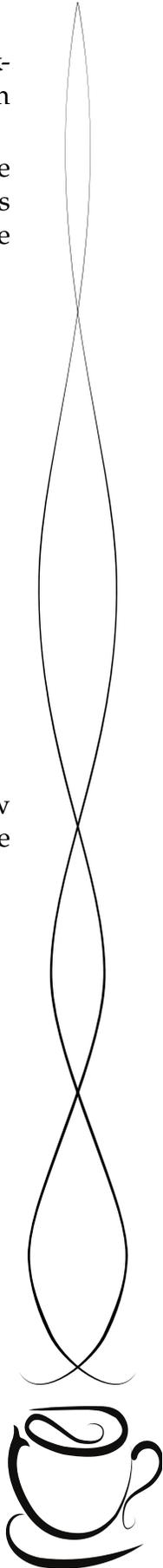
Here’s the breakdown by individual.

Rater	Correct Coffee	Correct Brand
Emma	2 (18%)	7 (64%)
Evan	0 (0%)	4 (36%)
Faye	1 (9%)	4 (36%)
Maciek	1 (9%)	4 (36%)
Spencer	3 (27%)	5 (45%)

Table 1: Proportion of correct coffee guesses made by each rater

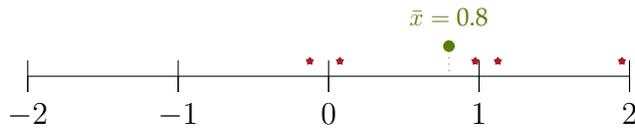
Detailed Results

The following pages contain detailed results for each coffee and each rater. Red stars show each rating (slightly offset so they don’t overlap). Green dots and the variable \bar{x} show the average rating. Fans of a coffee are those that rated it “Good” or “Very Good”.



Pilot Dangura

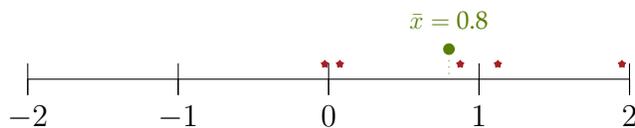
Ratings: Aroma



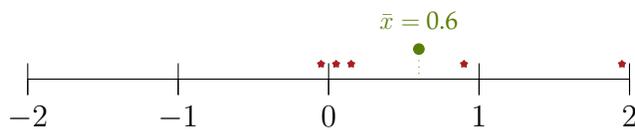
Fans of this brew:

Spencer
Evan
Emma

Ratings: Taste



Ratings: Aftertaste

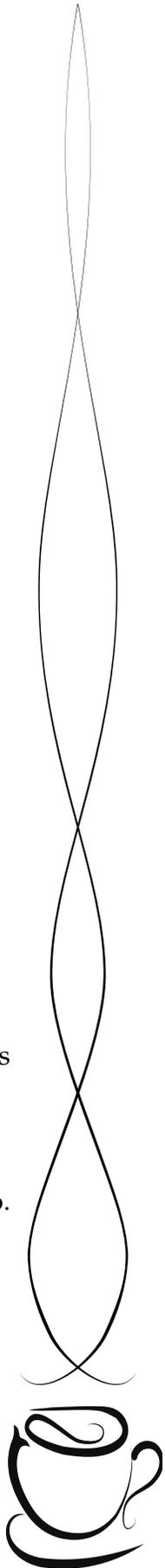


Deets:

Price: \$25/lb
Roast: NA
Pre-ground: FALSE

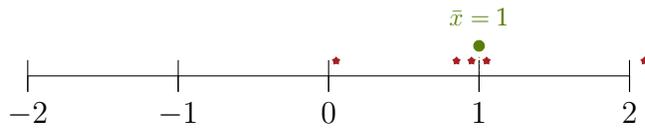
Comments

- Like s beautiful Italian woman, its full bodied and has a pleasant scent that lingers in the air.
-
- Smells nice, life normal coffee. Taste has a bit of kick, but not as much as the last two.
- Weak flavour, but not unpleasant.
- Spicy aggression



Pilot Tarime

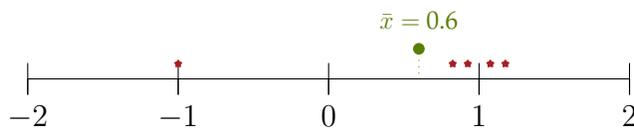
Ratings: Aroma



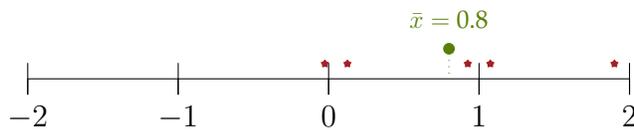
Fans of this brew:

Emma
Maciek
Spencer
Evan

Ratings: Taste



Ratings: Aftertaste

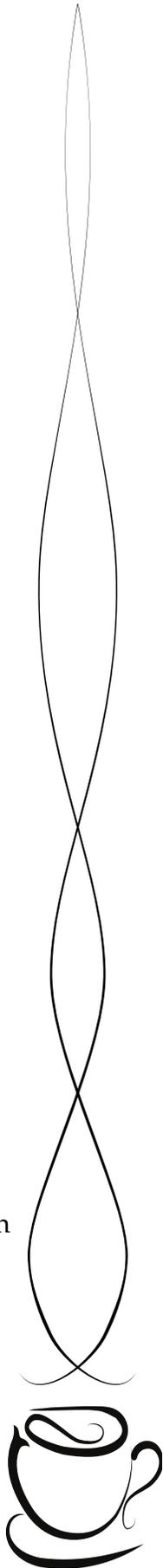


Deets:

Price: \$25/lb
Roast: NA
Pre-ground: FALSE

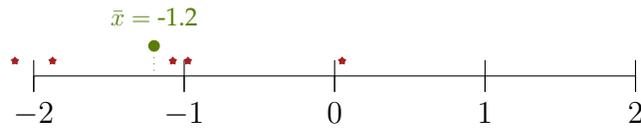
Comments

- Bold and beautiful
- Tangy, sharp, spicy.
- Wonderful! This person brewed a good cup of coffee.
- Very strong, different flavor - almost floral? But it the flower was punching you in the face. Not sure what's happening here. Don't know how to feel about this.
- kinda sour? dunno why

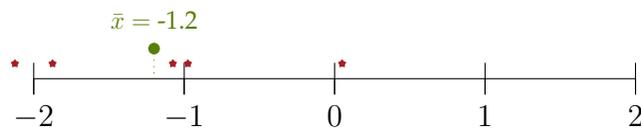


Faye's Mum's decaf

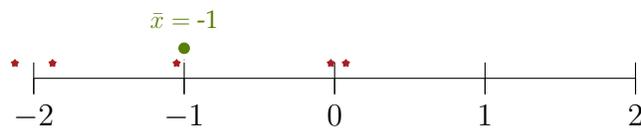
Ratings: Aroma



Ratings: Taste



Ratings: Aftertaste



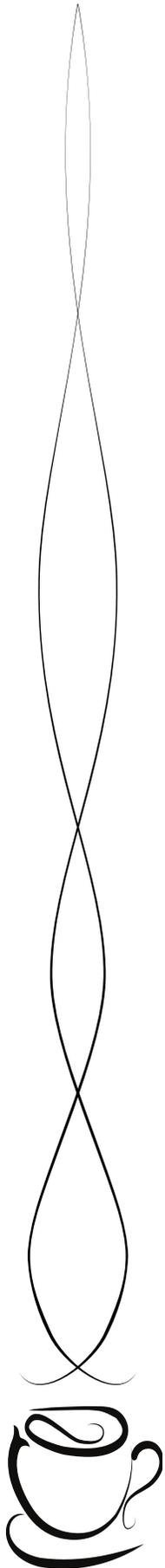
Comments

- No aroma, should be more like j
-
- Fuck you Macdonalds.
- Sweet-ish, medium strength
- Almost no smell. Intense initial flavor, but fades.

Fans of this brew:

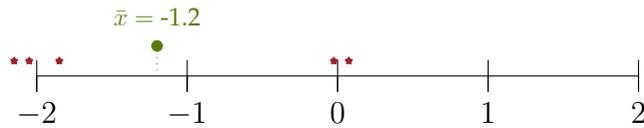
Deets:

Price: \$16/lb
 Roast: Decaf
 Pre-ground: TRUE



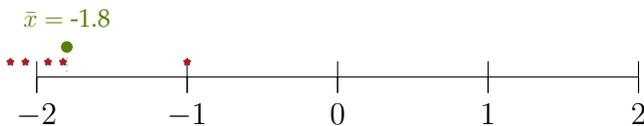
Kimbo Tradizione

Ratings: Aroma

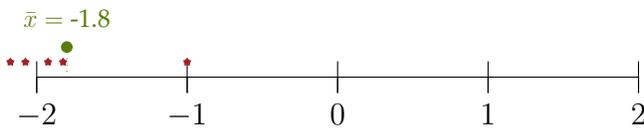


Fans of this brew:

Ratings: Taste



Ratings: Aftertaste

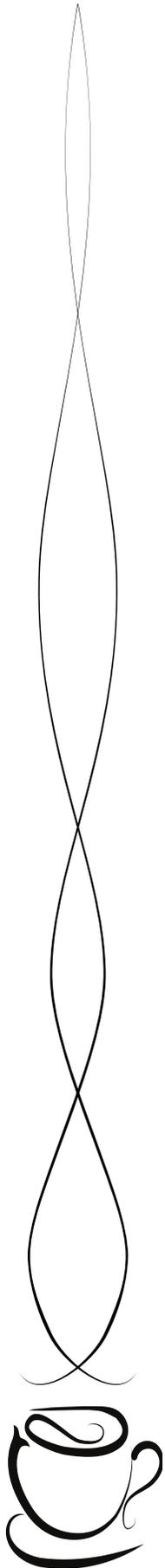


Deets:

Price: \$15/lb
 Roast: Medium-Dark
 Pre-ground: TRUE

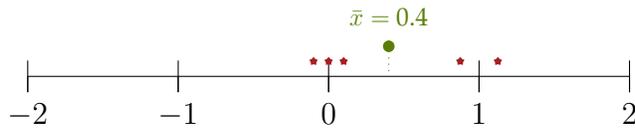
Comments

- Tastes like watery garbage
- No no no no
- the worst
- I would rather same the coffee that a cat shit out.
- Very high acidity. Lingers for a bit then fades. Taste is basically just acid though.



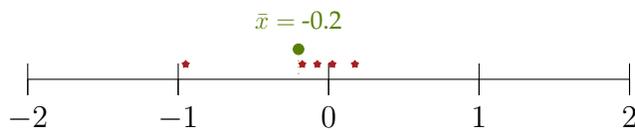
McCafe Medium Dark

Ratings: Aroma

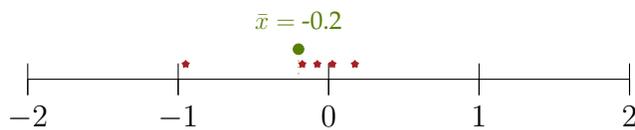


Fans of this brew:

Ratings: Taste



Ratings: Aftertaste

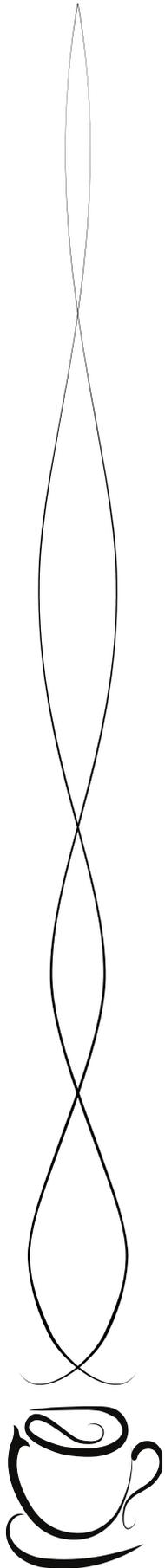


Deets:

Price: \$12/lb
 Roast: Medium-Dark
 Pre-ground: TRUE

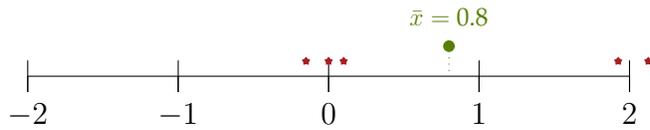
Comments

- You Sir Pilot make an acceptable coffee, that we should give to inmates
- Average joe
-
- The normal-est tasting coffee. Not sure what else to say.
- Almost zero flavour



Pilot La Palmas

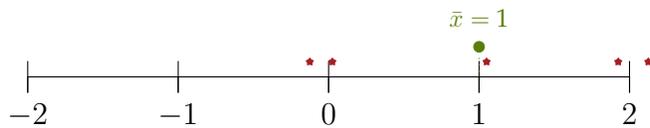
Ratings: Aroma



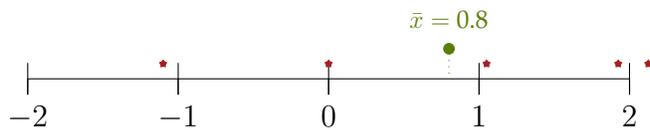
Fans of this brew:

Emma
Evan
Maciek

Ratings: Taste



Ratings: Aftertaste

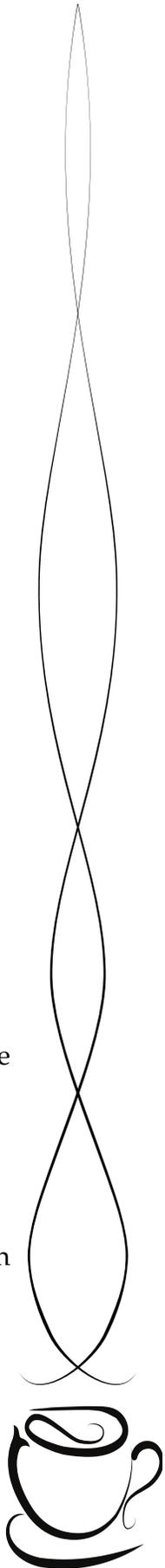


Deets:

Price: \$25/lb
Roast: NA
Pre-ground: FALSE

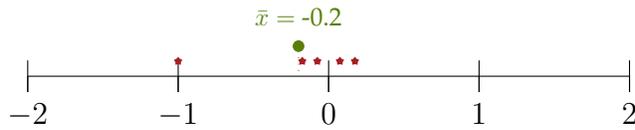
Comments

- Pilot you could work on improving your hopper to change this into a drinkable coffee rather than one I would pour on your lap.
- Afternoon delight
- actually good
- Has a slight tangy flavor, but very slight. Otherwise not that flavourful. Which makes it reasonably inoffensive.
- Little tangy, gentle flavour but good.



Kicking Horse Kick Ass

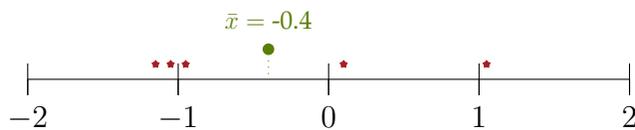
Ratings: Aroma



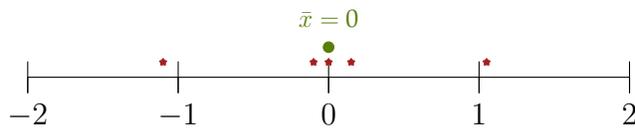
Fans of this brew:

Maciek

Ratings: Taste



Ratings: Aftertaste



Deets:

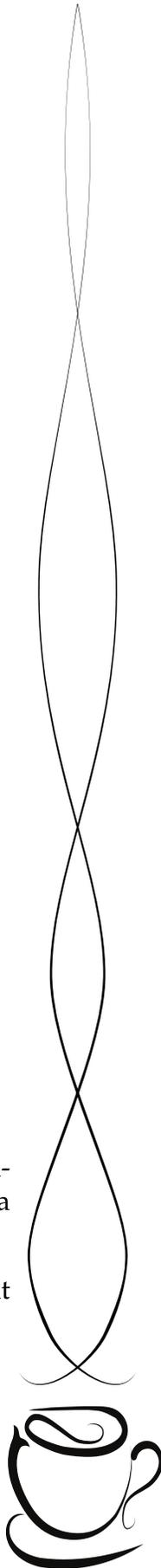
Price: \$16/lb

Roast: Dark

Pre-ground: FALSE

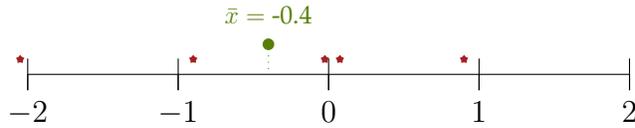
Comments

- Sweet and pleasant
- Average goat
- This barista decided to combine dishwasher with an unripe bean, to develop a unpleasant flavor that must have come from hippies in bc, that wanted to develop a weak ass coffee
- Bitter with not much additional flavor. Taste disappears from your mouth without lingering. Very plain.
- i mean water tastes good too



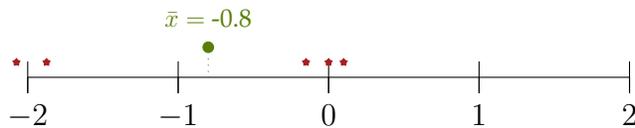
Blackfeather Blend

Ratings: Aroma



Fans of this brew:

Ratings: Taste



Ratings: Aftertaste

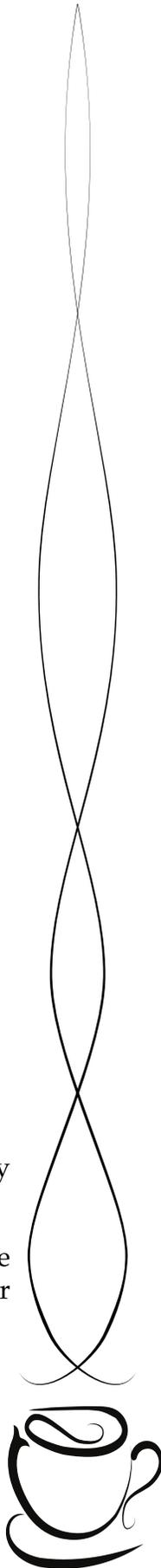


Deets:

Price: \$11/lb
 Roast: NA
 Pre-ground: TRUE

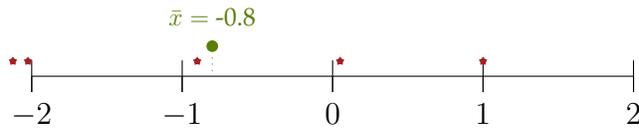
Comments

- Strong, acceptable but not great
- Yuck shit
- Smells almost smokey. Interesting flavor - definitely some flavour addition. Very bitter, with lingering bitterness.
- Oh! Someone understand the berry side of coffee, good for you! Its do bad you are still dirty hippies from bc and washed your feet in the pot that you used to roast your coffee.
- dark charcoal mess



Kicking Horse Three Sisters

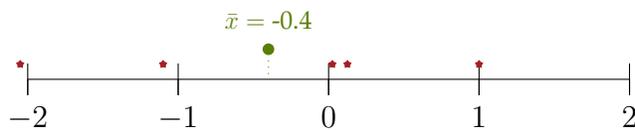
Ratings: Aroma



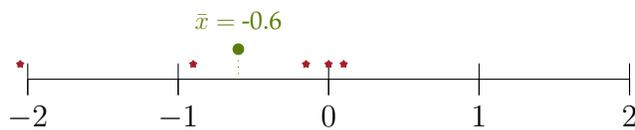
Fans of this brew:

Faye

Ratings: Taste



Ratings: Aftertaste



Deets:

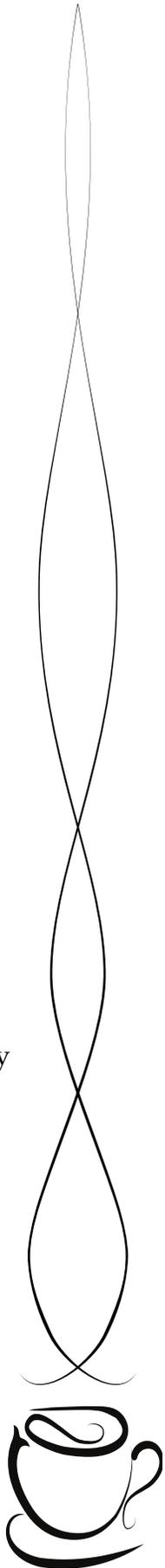
Price: \$16/lb

Roast: Medium

Pre-ground: FALSE

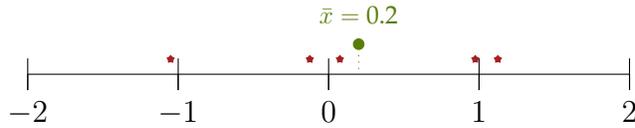
Comments

- Very neutral. Run of the mill coffee that you might get at any coffee shop. Not very bitter.
- bleh
- Did you put beans in or is this just water that has gone through the machine...
- A pleasant, gentle bitterness. Pretty drinkable.
- Needs milk



Kicking Horse Hola

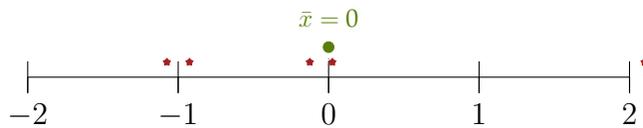
Ratings: Aroma



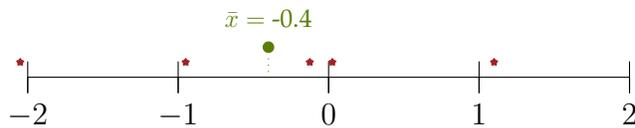
Fans of this brew:

Emma

Ratings: Taste



Ratings: Aftertaste



Deets:

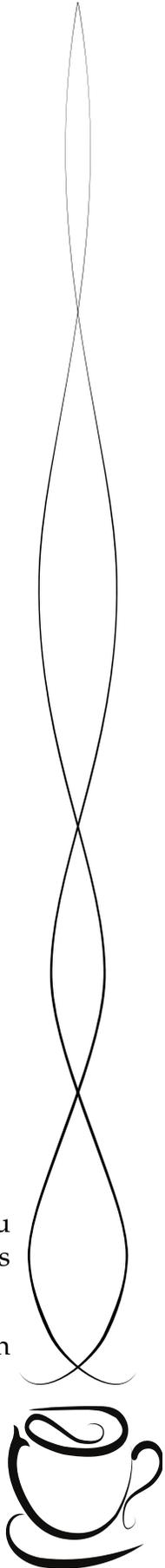
Price: \$16/lb

Roast: Light

Pre-ground: FALSE

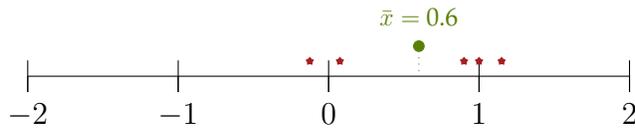
Comments

- Finely ground
- Little bit nasty tasting, the wrong kind of bitter.
- tastes like toilet water
- Oh? Did someone understand basic heating principals of a coffee bean, good for you get a gold star that we give children out for 'participating' in a soccer match that has no score.
- A bit more bite to it. Smells nice. Lingers a little longer. A little longer here with you. Bitter first taste, but the bitterness fades.



Kicking Horse Decaf

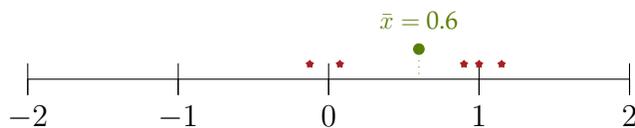
Ratings: Aroma



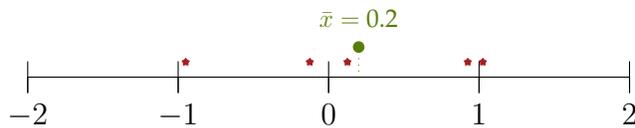
Fans of this brew:

Emma
Faye
Evan

Ratings: Taste



Ratings: Aftertaste

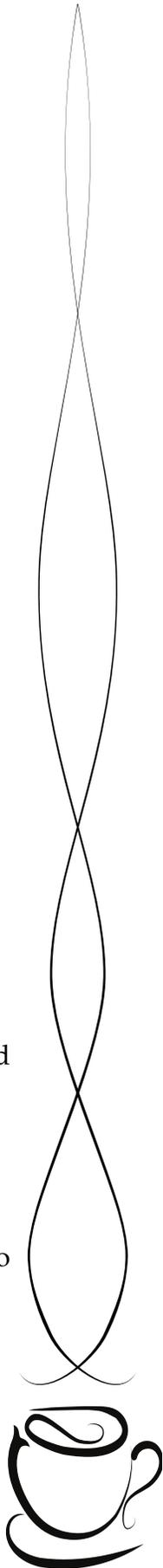


Deets:

Price: \$16/lb
Roast: Decaf
Pre-ground: FALSE

Comments

- A safe choice in coffee, if you want to live life safely than you get out my sight and never have your shame shadow darken a doorstep of light and awesomeness
- Weak, uneventful.
- Dark and brewing
- Actually pretty good. Maybe because it's slightly warmer than the others. Not too strongly flavored. Subtle notes.
-



Rater: Spencer

Coffee ordering:

Pilot Tarime (1)
Pilot Dangura (1)
Kicking Horse Decaf (0)
Kicking Horse Hola (0)
Blackfeather Blend (0)
Pilot La Palmas (0)
McCafe Medium Dark (0)
Kicking Horse Three Sisters (-1)
Kicking Horse Kick Ass (-1)
Kimbo Tradizione (-2)
Faye's Mum's decaf (-2)

Preferences:

Price(/\$10): 0.6

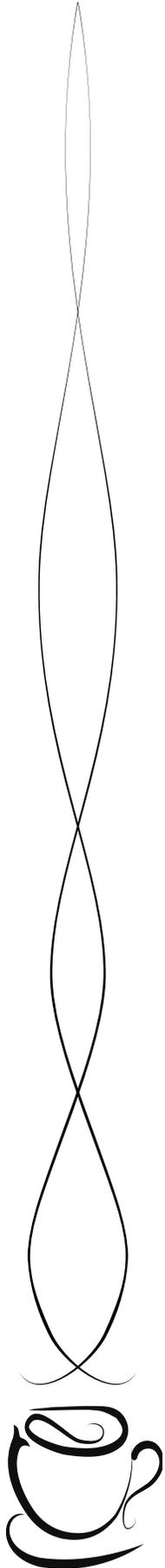
Fresh-ground: 0.6

A \$10 increase in coffee price causes this rater's Taste score to change by: **0.6**

Having fresh ground coffee (after statistically controlling for differences in price) causes this rater's Taste score to change by **0.6**

Correctly Gussed:

Pilot La Palmas
Blackfeather Blend
Kicking Horse Three Sisters



Rater: Evan

Coffee ordering:

Pilot La Palmas (2)
Pilot Dangura (2)
Kicking Horse Decaf (1)
Pilot Tarime (1)
McCafe Medium Dark (0)
Kicking Horse Hola (-1)
Kicking Horse Kick Ass (-1)
Faye's Mum's decaf (-1)
Kicking Horse Three Sisters (-2)
Blackfeather Blend (-2)
Kimbo Tradizione (-2)

Preferences:

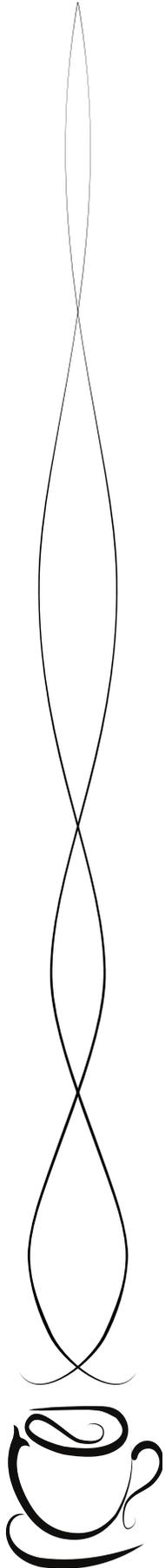
Price(/\$10): 2.4

Fresh-ground: 0

A \$10 increase in coffee price causes this rater's Taste score to change by: **2.4**

Having fresh ground coffee (after statistically controlling for differences in price) causes this rater's Taste score to change by **0**

Correctly Gussed:



Rater: Faye

Preferences:

Price(/\$10): -0.8

Fresh-ground: 1

A \$10 increase in coffee price causes this rater's Taste score to change by: **-0.8**

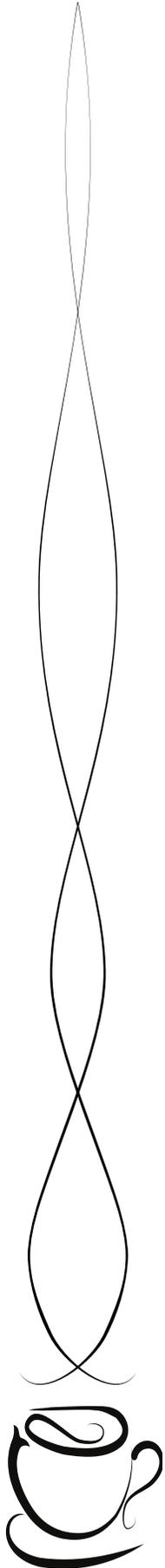
Having fresh ground coffee (after statistically controlling for differences in price) causes this rater's Taste score to change by **1**

Coffee ordering:

Kicking Horse Decaf (1)
Kicking Horse Three Sisters (1)
Kicking Horse Hola (0)
Blackfeather Blend (0)
Pilot La Palmas (0)
McCafe Medium Dark (0)
Pilot Dangura (0)
Kicking Horse Kick Ass (-1)
Kimbo Tradizione (-1)
Faye's Mum's decaf (-1)
Pilot Tarime (-1)

Correctly Gussed:

McCafe Medium Dark



Rater: Maciek

Preferences:

Price(/\$10): 0.6

Fresh-ground: 0.7

A \$10 increase in coffee price causes this rater's Taste score to change by: 0.6

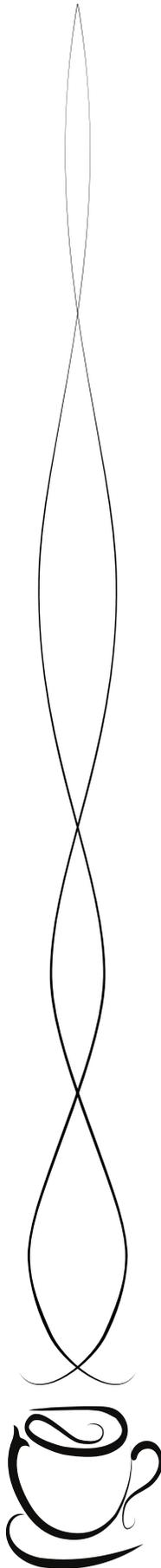
Having fresh ground coffee (after statistically controlling for differences in price) causes this rater's Taste score to change by 0.7

Coffee ordering:

Kicking Horse Kick Ass (1)
Pilot La Palmas (1)
Pilot Tarime (1)
Kicking Horse Decaf (0)
Kicking Horse Three Sisters (0)
Blackfeather Blend (0)
Faye's Mum's decaf (0)
Pilot Dangura (0)
Kicking Horse Hola (-1)
McCafe Medium Dark (-1)
Kimbo Tradizione (-2)

Correctly Gussed:

Pilot La Palmas



Rater: Emma

Coffee ordering:

Kicking Horse Hola (2)
Pilot La Palmas (2)
Kicking Horse Decaf (1)
Pilot Tarime (1)
Pilot Dangura (1)
Kicking Horse Three Sisters (0)
Kicking Horse Kick Ass (0)
McCafe Medium Dark (0)
Blackfeather Blend (-2)
Kimbo Tradizione (-2)
Faye's Mum's decaf (-2)

Preferences:

Price(/\$10): 0.4

Fresh-ground: 2.3

A \$10 increase in coffee price causes this rater's Taste score to change by: **0.4**

Having fresh ground coffee (after statistically controlling for differences in price) causes this rater's Taste score to change by **2.3**

Correctly Gussed:

Faye's Mum's decaf
Blackfeather Blend

