Agenda

1. Multiple regression with two quantitative variables

Warmup: Italian Restaurants (cont'd) Recall the Italian Restaurants from last time. Let's build a model now for the *Price* as a function of the *Food* rating and the location relative to 5th Avenue.

```
require(mosaic)
NYC <- read.csv("http://www.math.smith.edu/~bbaumer/mth241/nyc.csv")</pre>
qplot(data = NYC, x = jitter(Food), y = Price) + geom_smooth(method = "lm",
   se = 0)
mod.fe <- lm(Price ~ Food + East, data = NYC)</pre>
mod.fe
##
## Call:
## lm(formula = Price ~ Food + East, data = NYC)
##
## Coefficients:
## (Intercept)
                       Food
                                     East
## -17.430
                      2.875
                                    1.459
```

In-Class Activity

- 1. Interpret the coefficients of this model. What is the value of being on the East Side of Fifth Avenue?
- 2. Calculate the expected *Price* of a restaurant in the East Village with a *Food* rating of 23.
- 3. Use plotModel() to visualize your model in the data space.

```
plotModel(mod.fe, xlab = "Jittered food rating", ylab = "Average Price (US$)",
    system = "ggplot2")
```

Multiple Regression with a Second Quantitative Variable If X_2 is a quantitative variable, then we have

$$\hat{Y} = \hat{\beta}_0 + \hat{\beta}_1 \cdot X_1 + \hat{\beta}_2 \cdot X_2$$

Notice that our model is no longer a line, rather it is a *plane* that lives in three dimensions!

Example: Italian Restaurants (continued) Now suppose that we want to improve our model by considering not only the quality of the *Food*, but also the quality of the *Service*. We can do this in **R** by simply adding another variable to our regression model.

mod.fs <- lm(Price ~ Food + Service, data = NYC)
coef(mod.fs)
(Intercept) Food Service
-21.158582 1.495369 1.704101</pre>

- 1. Interpret the coefficients of this model. What does the coefficient of *Food* mean in the real-world context of the problem? *Service*?
- 2. How important is *Service* relative to *Food*? Is it fair to compare the two coefficients?
- 3. Use makeFun() to find the expected *Price* of a restaurant with a *Food* rating of 21 but a *Service* rating of 28.
- 4. Calculate the residual for San Pietro. Is it overpriced?

```
filter(NYC, Restaurant == "San Pietro")
## Case Restaurant Price Food Decor Service East
## 1 44 San Pietro 58 24 21 23 1
```

5. What if we added all three explanatory variables to the model? What geometric shape would we have then?