1 Data Prep, EDA, and Theory development

1.1 Varaible Selection:

Generally helpful variables Number ofbedrooms and bathrooms overall size years since remodeling years since build how many spots in the garage

ALSO! If a house was sold multiple times, this would imply a cluster by house; or possibly by area as well; so this is interesint wrt the standard error of the eestimated parameters !!!

relevant variables: SalePrice

class of the building MSSubClass, BldgType —¿ eg one family home etc,

area in city/ district MSZoning, Neighborhood: Physical locations within Ames city limits

size of the property LotArea –¿ size of the overall lot/ property,

overall quality contains the overall quality of the house and the condition together IMPORTANT! Is the degree of finish of the house here/ i mean: is the house really finished and is this reflected in this variable? OverallQual – overall quality, OverallCond –¿ overall condition of the house,

age and need or remodeling of he house inpsulateded here YearBuilt, YearRemodAdd,

heating overall HeatingQC CentralAir Electrical

size of the house 1stFlrSF: First Floor square feet 2ndFlrSF: Second floor square feet

nuber of bathrooms BsmtFullBath: Basement full bathrooms BsmtHalfBath: Basement half bathrooms FullBath: Full bathrooms above grade HalfBath: Half baths above grade

Bedroom: Number of bedrooms above basement level Kitchen: Number of kitchens

Functional: Home functionality rating

GarageCars: Size of garage in car capacity GarageArea: Size of garage in square feet

sales features MoSold: Month Sold YrSold: Year Sold SaleType: Type of sale SaleCondition: Condition of sale