|  |  |
| --- | --- |
|  | **Work Integrated Learning Programmes Division**  **M.Tech (Data Science and Engineering)**  https://lh6.googleusercontent.com/ENp5iVIZzIHv2b4n4mkS4vHt4UkEBlWYmz738CY-LLW4hqPglKHQ6jUh7UR0d4Ymmrr709Maa1nSxWlKGCsAkajArO1QhFOMQzxwHhEhk0zqlMRi4_H7oj2RIxwzkiz_qKiY_Giy |

**Data Visualization & Interpretation   
(DSECLZG555)**

**Assignment 1 – PS7 - [CONSTRUCTION DATA ANALYSIS]**

**Group:**

DVI GROUP 147

**Team Members:**

Ashwini Bangalore Ramesh

Rahul Paul

Ashutosh Kumar

**Contents**

[1. **Know Your Audience** 3](#_Toc75977579)

[**2.** **What ?** 4](#_Toc75977580)

* [**Data Preparation** 4](#_Toc75977581)

[**3.** **Present the BIG IDEA** 4](#_Toc75977582)

[**4.** **HOW ?** 5](#_Toc75977583)

[a) Chart 1: 5](#_Toc75977584)

[b) Chart 2: 5](#_Toc75977585)

[c) Chart 3: 6](#_Toc75977586)

[d) Gestalt Principle Employed 6](#_Toc75977587)

[e) Pre-attentive attributes 6](#_Toc75977588)

[**5.** **Dashboard** 7](#_Toc75977589)

# **Know Your Audience**

1. List the primary groups or individuals to whom you’ll be communicating.

**»» The Department of Housing and Urban Development**

1. If you had to narrow that to a single person, who would that be?

**»» Secretary of Department of Housing and Urban Development**

1. What does your audience care about?

**»» High level  
 The Department of Housing and Urban Development uses the estimates to develop and   
 evaluate housing programs.**

**Granular level   
 1. Correct estimate of construction value of New Single-Family Houses Under Construction   
 through both constant Quality (Laspeyres) and Price Deflator (Fisher).**

**2. Correct estimate of Sell price of New Single-Family Houses in whole US and specific region   
 wise like NE , MW , SO and WE.**

1. What action does your audience need to take?

**»» Correct reporting of Construction value and average sell price of new Single family houses   
 in current year compared to previous year. It depends on understanding which index is   
 appropriate.**

1. What is at stake? What is the benefit if the audience acts in the way you want them to? What are   
    the risks if they don’t?

**»» *At stake:*  
 The Department of Housing and Urban Development uses the estimates to develop   
 and evaluate housing programs.**

***Benefit:***

**1. Able to find out which price index is suitable (Under construction Laspeyres or Fisher)**

**2. Able to take appropriate decision based on trend of under construction and sell   
 price index spread over the years to target region (NE,MW,SO,WE)for housing program.**

***Risks:*  
 Decision of target region and price index calculation could go wrong leading to ineffective   
 housing programs.**

# **What ?**

* What is that you are trying to communicate? What questions are you trying to answer/display in   
   your visualizations? Write these as specific questions. You need to come up with 3 questions at   
   least, each of which will be answered using one Viz.

**»» 1.Which price index (Laspeyres or Fisher) is better?**

**2. Which region has better sell price opportunity (US vs NE, MW, SO and WE)?**

**3.** **How accurately are the current and 2005 houses sold percentage changes are predicted   
 by the Laspeyers price index calculations?**

## **Data Preparation**

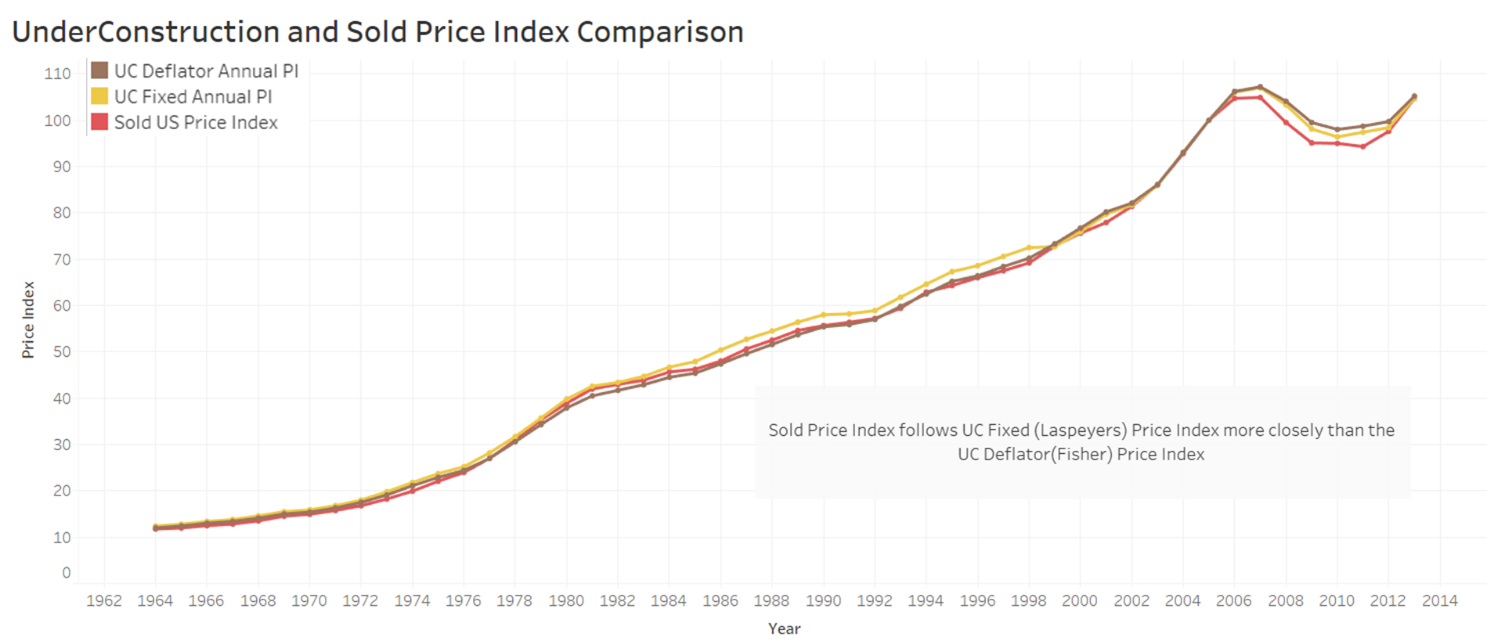
* **Merged multiple headers to Single cell header and headers are renamed accordingly.**
* **Sold Price index tab is unchanged and other 3 tabs having sold price are merged.**
* **Null values are ignored.**
* **Created calculated fields in Tableau to visualize the trending of house %change difference   
   against Price Index difference**

# **Present the BIG IDEA**

**Laspeyers price indexing gives better calculation for under construction and sold price index based on geographic location and also consider inflation. So we recommend to use laspeyers price indexing going forward.**

# **HOW ?**

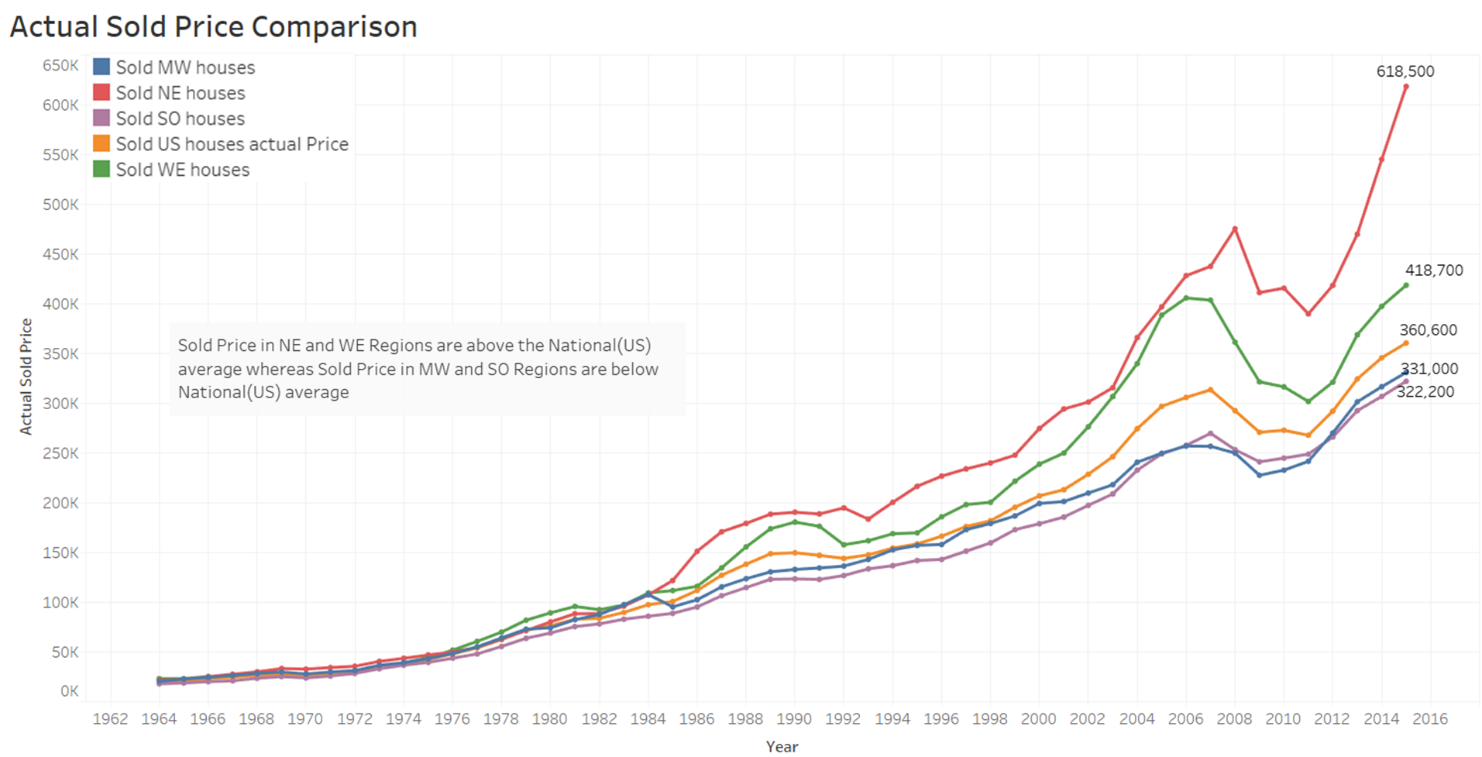
## Chart 1:



**What:** Line chart is used to compare sold price index of with that Laspeyers and Fisher index.

**Why:** The line chart is used to visualize trend in data over intervals of time (Yearly) and as smaller changes exist, line chart are better.

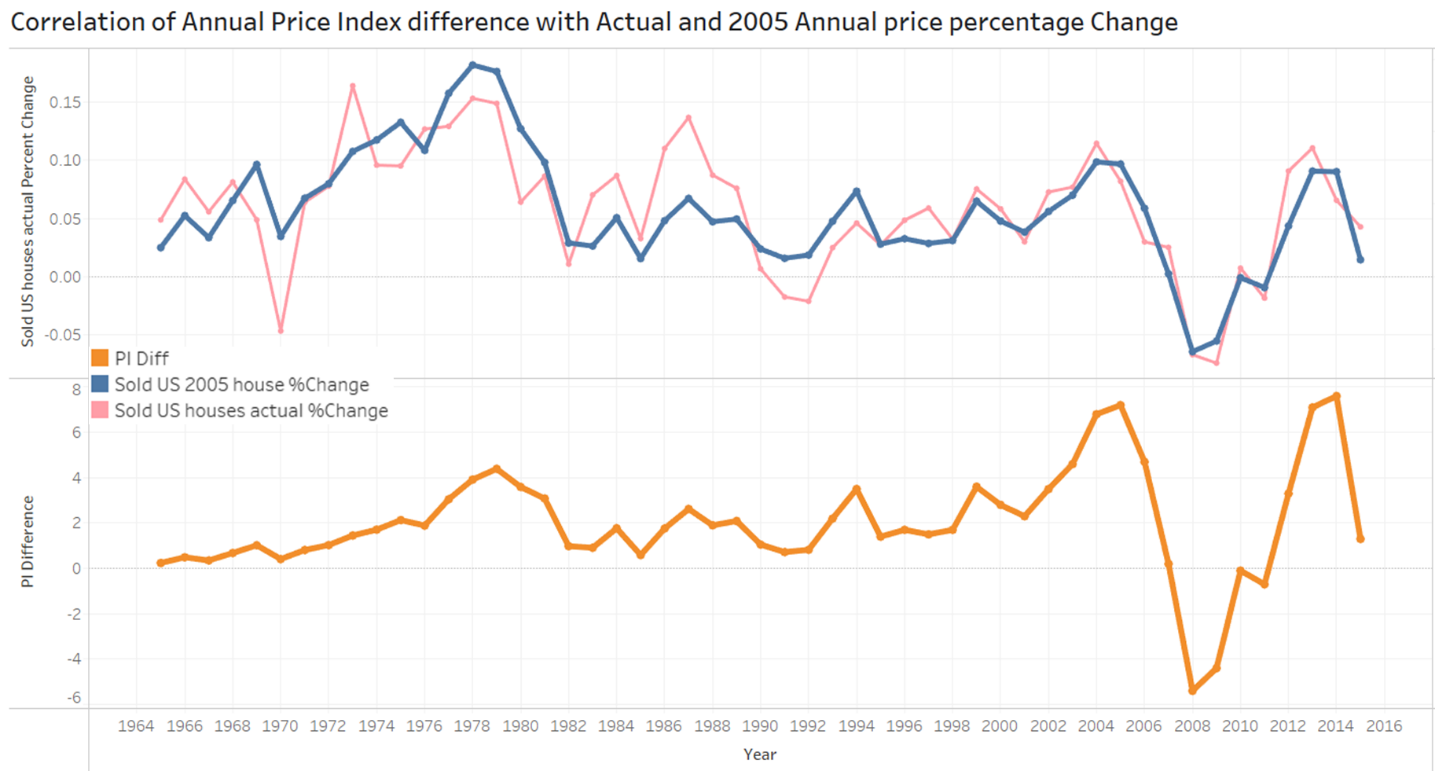
## Chart 2:



**What:** Line chart is used to compare region wise sold price index across years.

**Why:** The line chart is used to visualize trend in data over intervals of time (Yearly) and as smaller changes exist, line chart are better.

## Chart 3:



**What:** Line chart is used to find the correlation of price index with actual sold & 2005 sold annual price percentage

**Why:** The line chart is used to visualize trend in data over intervals of time (Yearly) and as smaller changes exist, line chart are better.

## Gestalt Principle Employed

1. Gestalt’s Law of Similarity – Colour line to connect data points
2. Gestalt’s Law of Symmetry – Chart 3 showing trend symmetry
3. Gestalt’s Law of Proximity – Closeness of data labels and data markers
4. Gestalt’s Law of Connection – Overall trend
5. Gestalt’s Law of Closure – No border around the graph as well as individual text boxes.

## Pre-attentive attributes

* All charts make use of color to differentiate and highlight the important aspect of the charts.   
   The non-important aspects are greyed out.
* Most of the charts make use of same axis to maintain context among them so that the viewer   
   does not get false sense of scale. This is not followed only for chart 3 which requires an   
   individual axis to be able to show the value.
* In chart 3, thicker lines are used to visualize the trending between the house percentage   
   change difference and price index difference.

# **Dashboard**

