

Marketing Applications of Data Mining Analysis Team Project

The Blue Notes

Michael Casserly

Francesca DeLisi

Ryosuke Okumura

Hyok-Hee Ho

Andrew Weisselberg

Executive Summary

■ Project

- Use customer databases of hospital cancer patients to identify relevant demographic information for use in target marketing
 - Utilize demographic variables in regional data set – NY, NJ, CT

■ Product

- A chemotherapy treatment that reduces the side effects of cancer, specifically nausea and vomiting

■ Market Potential

- 1.4 million new cancer patients every year in the United States

Identifying Market Segments

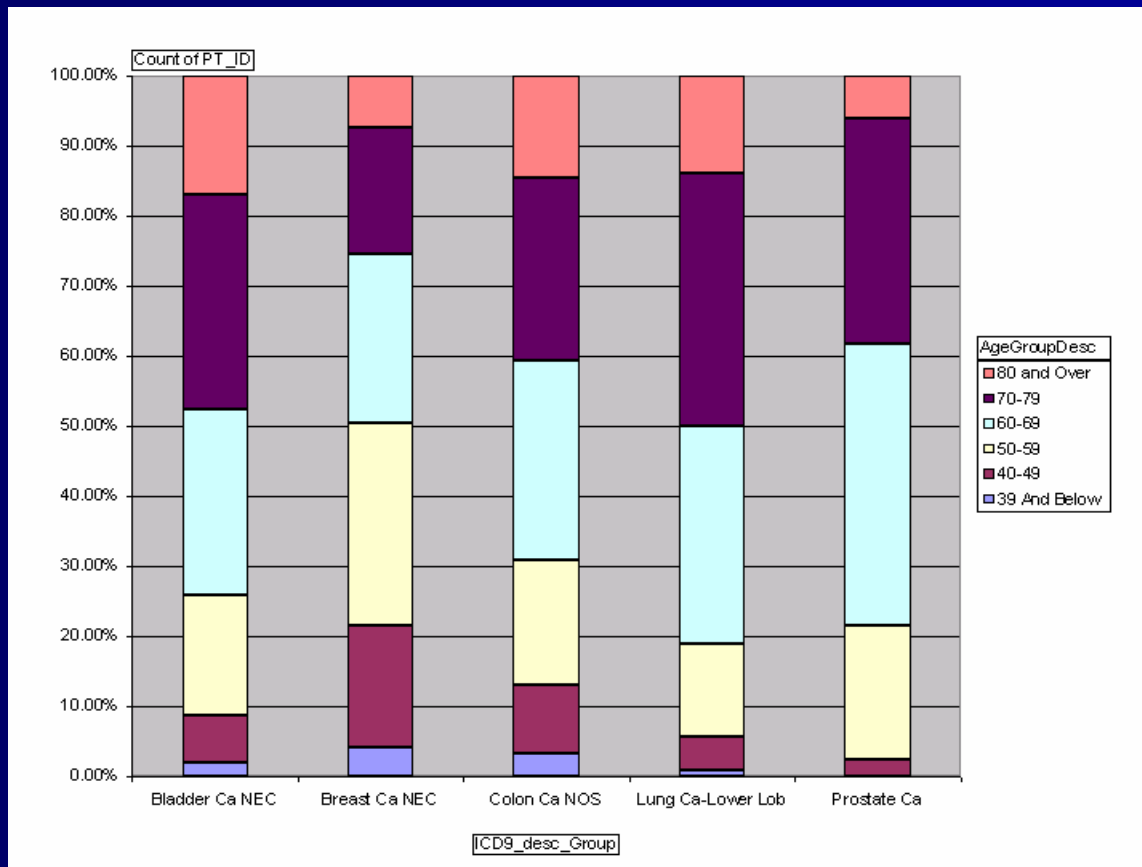
- Data mining task
 - Classification
- Determine relationships between cancer and independent variables (age, income, marital status, race, sex)
- Data Transformation

Findings

■ Age Demographics

- Largest % of male patients falls into 60 - 80 (66%)
 - Distribution is similar for most of the cancers, prostate cancer exception
- Largest % of female patients falls into the 50-80 (71%)
 - Majority of Breast Cancer patients are between 50 and 70

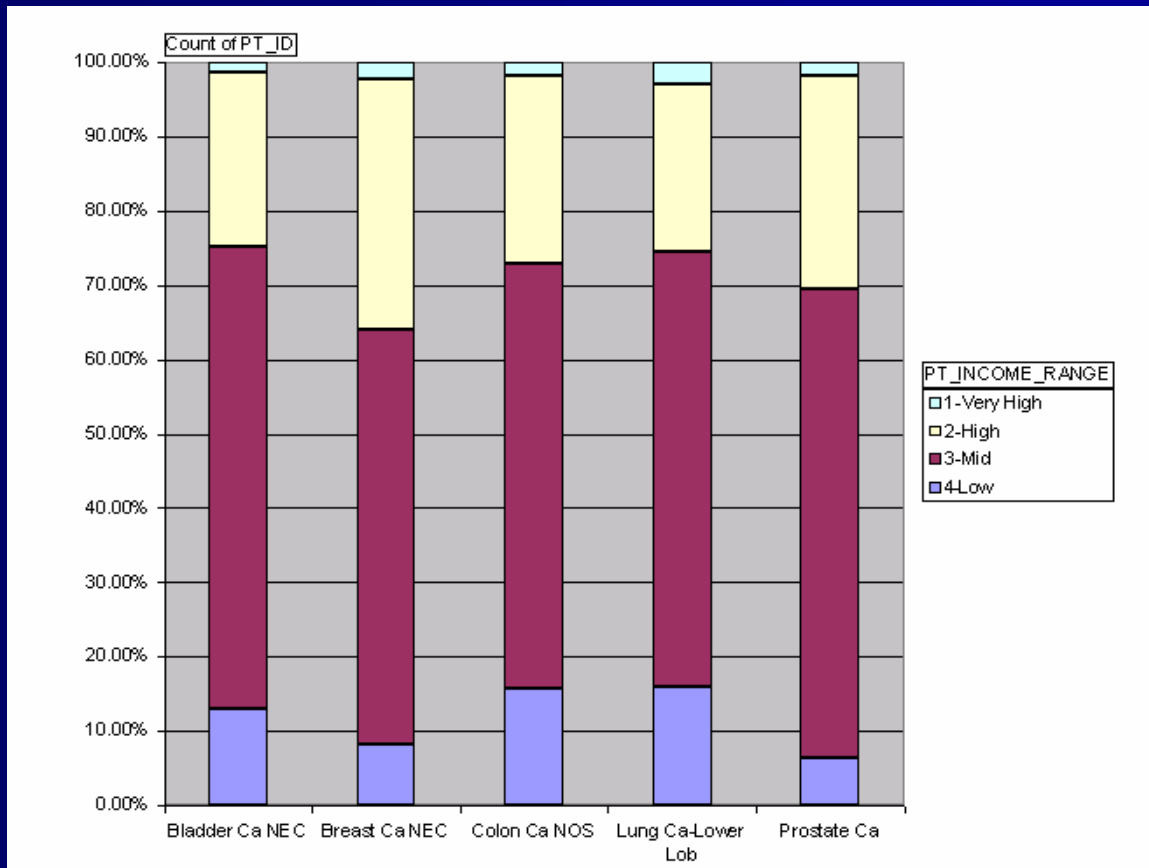
Age Demographic Chart



Findings

- Income Demographics
 - Largest % of male and female patients fall into the middle income bracket, \$31,000 – \$75,000 (59%)
 - Income distribution consistent among all cancers

Income Demographic Chart



Findings

■ Race Demographics

- Largest % of patients are Caucasian (85%)
- Second largest % are Black Non Hispanic (6.58%)

Conclusions

- Based on the extrapolated demographic information we have a clear and static target market that represents a significant economic opportunity
- We therefore request \$500,000 in incremental funding to secure additional patient databases in order to validate our findings on a national basis