database design project
ieor 115: final presentation

PROJECT: JUICE
Local + Cold-Pressed + Raw

team #6
Alicia Zecherle
Ami Bigit
Ben Lewis
Khoa Tran
Kimberly Lam
Mana Nabati
Morgan Fabian
Steven Belding
our client: Rachel Malsin
founder of PROJECT: JUICE
relational schema

project overview

1. Person(PersonID, Lname, Fname, MI, Street, City, State, Zip, Telephone, Email)
2. Customer(CID¹, Customer_Type, Group_Name, Bdate, Gender, Company, County, Tax_Rate, Note)
3. Employee(EID¹, SSN, Bdate, Job_Title, Salary, Hourly_Wage)
4. Facility_Owner(FOID¹, Company_Name)
5. Potential_Customer(ContactID¹, Place_of_Contact, Date_Of_Contact, Company_Name, SamplingEvent²⁰)
6. Vendor_Contact(ContactID¹, Job_Title, Company_Name)
7. Investor(InvID¹, Company_Name)
8. Other_Person(ContactID¹, Description)
9. Order(OID, Date_Placed, Time_Placed, CID², Input_Type, TID³¹, Received_By³)
10. Facility(FacilityID, Name, Street, City, State, Zip, Telephone, Facility_Type)
11. Vendor(VID, Company_Name, Type, ContactID⁶, Street, City, Zip, Telephone, Email)
12. Product(PID, Pname, Season)
13. Ingredient(PID¹², Season)
14. Juice(PID¹², Shelf_Life)
15. Tea(PID¹², Shelf_Life)
16. Bottle(PID¹², Size, Material, Cost, VendorID¹¹)
17. Label(PID¹², Size, Printed_At, For_Bottle_Size, Cost, VendorID¹¹)
18. Bottle_Carriers(PID¹², Carrier_Type, Color, Cost, Insulated)
19. Cleanse(PID¹², Cleanse_Name, Number_of_Days, Price)
20. Sampling_Event(EventID, Date, Place, Event_Name, Type, Event_Description)
21. **Product_Record_Sellable** (PID\(^{12}\), Date, Instance, Mixed_BID\(^{10}\), Mixed_BID\(^{24}\), Bottle\(^{16}\), Label_Type\(^{17}\), Shipment\(^{25}\), VID\(^{11}\), Price)
22. **Product_Record_NonSellable** (PID\(^{13}\), Date_Bought, Instance, Weight, VID\(^{11}\), Cost, TID\(^{30}\))
23. **Single_Ingredient_Batch** (SI_BID, PID\(^{22}\), Date_Bought\(^{22}\), Instance\(^{22}\), Ingredient\(^{13}\), Original_Weight, Produced_Volume, Date_Made, BagID\(^{27}\), BladeID\(^{27}\))
24. **Mixed_Batch** (Mixed_BID, BName, Date_Made, Date_Finalized, Date_Expires, PID\(^{12}\), Treatment, Treatment_Date, Final_Volume)
25. **Shipment** (SID, OID\(^{9}\), Date, Delivered_By\(^{3}\), FedEx_Option, Bottle_Carrier\(^{18}\), Delivered_From\(^{10}\))
26. **Non-Product** (NID, Name, Type, Description, Restock_Every)
27. **Equipment** (EquipID\(^{26}\), EName, Type)
28. **Office_Supplies** (SupplyID\(^{26}\), OName, Type)
29. **Non-Product_Record** (NID\(^{26}\), Date_Purchased, Instance, Location\(^{10}\), VID\(^{11}\))
30. **Transaction** (TID, Amount, TDate, TTime)
31. **Withdrawl** (TID\(^{30}\), Type, SalaryEmployeeID\(^{3}\), Facility_Rent\(^{4}\), Refund_Order\(^{9}\), Vendor_Purchase\(^{11}\), Amount)
32. **Deposit** (TID\(^{30}\), OID\(^{9}\), InvestorID\(^{7}\), Type, Amount)
33. **Updates_Order** (OID\(^{9}\), EID\(^{3}\), UDate, UTime)
34. **Employee_Makes_SIBatch_With_Equip** (EID\(^{3}\), SIBID\(^{23}\), EquipID\(^{27}\), Time_Started, Time_Ended)
35. **Employee_Makes_MixedBatch_With_Equip** (EID\(^{3}\), MixedBID\(^{24}\), EquipID\(^{27}\), Time_Started, Time_Ended)
36. **Employee_Makes_Sellable_Product** (EID\(^{3}\), PID\(^{21}\), PDate\(^{21}\), PInstance\(^{21}\))
37. **Order_Details** (OID\(^{9}\), PID\(^{12}\), Quantity)
38. **Vendor_Contact_List** (VID\(^{11}\), Vendor_Contact\(^{6}\))
relational schema

project overview

39. Product_Facility_To_Facility_Sellable(PID21, PDate21, Pinstance21, Original_Facility10, New_Facility10, TransportDate)
40. Product_Facility_To_Facility_Nonsellable(PID22, PDate22, Pinstance22, Original_Facility10, New_Facility10, TransportDate)
41. NonProduct_Facility_To_Facility(NID29, NPDate29, NPIInstance29, Original_Facility10, New_Facility10, TransportDate)
42. Vendor_Ships_To_Facility(VID11, Ships_To10, SDate, STime)
43. Vendor_Carries_Product(VID11, Product_ID12, Season)
44. Vendor_Carries_NonProduct(VID11, NonProduct_ID26)
45. Prod Includes_Prod(Product12, Sub_Product12, Liquid_Volume)
46. MBatch_Contains_SI_Batch(Mixed_BID24, SI_BID23, Liquid_Volume)
47. NonSellable_Makes_Batch(Nonsellable_PID22, NBDate22, NBInstance22, SI_BID23)
48. Facility_Owned_By(Facility_ID10, Facility_Owner4)
49. Ship_Is_Filled_By(SID25, PID21, Date21, Instance21)
50. Custom_Cleanse(PID19, OID9, Quantity, Juice114, Juice214, Juice314, Juice414, Juice514, Juice614, Juice714, Juice814, Juice914, Juice1014, Tea115)
51. Juices_At_Event(PID12, EventID20)
database implementation

how they will be using it:

- **FORMS**
  - Entering orders details
  - Entering and editing product history
  - Updating customers and contacts

- **REPORTS**
  - Lists of daily orders
  - Lists of orders by Zip Code for Tax Info
  - Lists of reoccurring customer orders

- **QUERIES:** Improving production, forecasting, and marketing strategy
database implementation:
forms and switchboard
database implementation
query 1: contaminated

WHAT INGREDIENTS?

batch why?

WHICH PRODUCTS?

LOTS OF VENDOR PRODUCE

JUICES

Customers

Complaint
Given a particular customer complaint, find ALL customers that may have also been affected
query 1: contaminated batch
lines 1-45

```
SELECT c.[Number of Possible Contaminated], *
FROM person AS p, [Query_4_#_Contam] AS c
WHERE p.PersonID in (;

select o.cid
from [order] o
where o.oid in (;

select s.oid
from shipment s
where s.sid in (;

select z.sid
from ship_is_filled_by z
inner join (;

select R.Instance, R.PID, R.pDate
from Product_Record_Sellable R
where R.Mixed_BID1 in (;

select B.mixed_BID
from mbatch_contains_sibatch B
where B.SI_BID in (;

select distinct BB.SI_BID
from mbatch_contains_sibatch bb
where BB.Mixed_BID in (;

select distinct RR.mixed_BID1
from Product_Record_Sellable RR
where RR.PID in (;

select zz.PID from ship_is_filled_by zz where zz.SID in (;
select ss.sid from shipment ss where ss.oid in (;
select oo.oid from [order] oo where oo.cid = 115753140 AND oo.Date_Placed = #1/2/2013#))
)
AND RR.pDate in (;
select zz.pDate from ship_is_filled_by zz where zz.SID in (;
select ss.sid from shipment ss where ss.oid in (;
select oo.oid from [order] oo where oo.cid = 115753140 AND oo.Date_Placed = #1/2/2013#))
)
AND RR.Instance in (;
select zz.Instance from ship_is_filled_by zz where zz.SID in (;
select ss.sid from shipment ss where ss.oid in (;
select oo.oid from [order] oo where oo.cid = 115753140 AND oo.Date_Placed = #1/2/2013#))
)
```
query 1: contaminated batch
lines 46-90

```sql
) OR BB.Mixed_BID in (  
    select distinct RR.mixed_BID2  
    from Product_Record_Sellable RR  
    where RR.PID in (  
        select zz.PID from ship_is_filled_by zz where zz.SID in (  
            select ss.sid from shipment ss where ss.oid in (  
                select oo.oid from [order] oo where oo.cid = 115753140 AND oo.Date_Placed = #1/2/2013#))  
        AND RR.pDate in (  
            select zz.pDate from ship_is_filled_by zz where zz.SID in (  
                select ss.sid from shipment ss where ss.oid in (  
                    select oo.oid from [order] oo where oo.cid = 115753140 AND oo.Date_Placed = #1/2/2013#))  
            AND RR.Instance in (  
                select zz.Instance from ship_is_filled_by zz where zz.SID in (  
                    select ss.sid from shipment ss where ss.oid in (  
                        select oo.oid from [order] oo where oo.cid = 115753140 AND oo.Date_Placed = #1/2/2013#))  
                    )  
            )  
        )  
    )  
) OR R.Mixed_BID2 in (  
    select B.mixed_BID  
    from mbatch_contains_sibatch B  
    where B.SI_BID in (  
        select distinct BB.SI_BID  
        from mbatch_contains_sibatch bb  
        where BB.Mixed_BID in (  
            select distinct RR.mixed_BID1  
            from Product_Record_Sellable RR  
            where RR.PID in (  
                select zz.PID from ship_is_filled_by zz where zz.SID in (  
                    select ss.sid from shipment ss where ss.oid in (  
                        select oo.oid from [order] oo where oo.cid = 115753140 AND oo.Date_Placed = #1/2/2013#)  
                    )  
                    AND RR.pDate in (  
                        select zz.pDate from ship_is_filled_by zz where zz.SID in (  
                            select ss.sid from shipment ss where ss.oid in (  
                                select oo.oid from [order] oo where oo.cid = 115753140 AND oo.Date_Placed = #1/2/2013#)  
                            )  
                        )  
                    )  
                )  
            )  
        )  
    )  
);
query 1: contaminated batch
lines 91-126

```sql
select oo.oid from [order] oo where oo.cid = 115753140 AND oo.Date_Placed = #1/2/2013#

AND RR.Instance in ( select zz.Instance from ship_is_filled_by zz where zz.SID in ( select ss.sid from shipment ss where ss.oid in ( select oo.oid from [order] oo where oo.cid = 115753140 AND oo.Date_Placed = #1/2/2013# ) ) )

) OR BB.Mixed_BID in ( select distinct RR.mixed_BID2 from Product_Record_Sellable RR where RR.PID in ( select zz.PID from ship_is_filled_by zz where zz.SID in ( select ss.sid from shipment ss where ss.oid in ( select oo.oid from [order] oo where oo.cid = 115753140 AND oo.Date_Placed = #1/2/2013# ) ) ) AND RR.pDate in ( select zz.pDate from ship_is_filled_by zz where zz.SID in ( select ss.sid from shipment ss where ss.oid in ( select oo.oid from [order] oo where oo.cid = 115753140 AND oo.Date_Placed = #1/2/2013# ) ) ) AND RR.Instance in ( select zz.Instance from ship_is_filled_by zz where zz.SID in ( select ss.sid from shipment ss where ss.oid in ( select oo.oid from [order] oo where oo.cid = 115753140 AND oo.Date_Placed = #1/2/2013# ) ) ) )

) as a on a.Instance = z.Instance and a.PID = z.PID and a.pDate = z.pDate

);
query 1: contaminated batch

Given a bad order, find ALL products that may have also been affected and the probability that they are also bad

DAMAGE CONTROL

Step 1: Find all possibly contaminated batches given a customer complaint.
query 1: contaminated batch

Step 2: Find all possibly contaminated single ingredient batches and customers with products containing those batches.

<table>
<thead>
<tr>
<th>SI_BID</th>
<th>PersonID</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>108959616</td>
</tr>
<tr>
<td>3</td>
<td>108959616</td>
</tr>
<tr>
<td>4</td>
<td>108959616</td>
</tr>
<tr>
<td>5</td>
<td>108959616</td>
</tr>
<tr>
<td>6</td>
<td>108959616</td>
</tr>
<tr>
<td>9</td>
<td>108959616</td>
</tr>
<tr>
<td>10</td>
<td>108959616</td>
</tr>
<tr>
<td>100</td>
<td>108959616</td>
</tr>
<tr>
<td>1</td>
<td>114543430</td>
</tr>
<tr>
<td>2</td>
<td>114543430</td>
</tr>
<tr>
<td>3</td>
<td>114543430</td>
</tr>
<tr>
<td>4</td>
<td>114543430</td>
</tr>
<tr>
<td>5</td>
<td>114543430</td>
</tr>
<tr>
<td>6</td>
<td>114543430</td>
</tr>
<tr>
<td>8</td>
<td>114543430</td>
</tr>
<tr>
<td>9</td>
<td>114543430</td>
</tr>
<tr>
<td>10</td>
<td>114543430</td>
</tr>
<tr>
<td>11</td>
<td>114543430</td>
</tr>
<tr>
<td>100</td>
<td>114543430</td>
</tr>
<tr>
<td>1</td>
<td>115753140</td>
</tr>
<tr>
<td>2</td>
<td>115753140</td>
</tr>
<tr>
<td>3</td>
<td>115753140</td>
</tr>
<tr>
<td>4</td>
<td>115753140</td>
</tr>
<tr>
<td>10</td>
<td>115753140</td>
</tr>
</tbody>
</table>
query 1: contaminated batch

Step 3: Determine which customers to contact and find the probability that they have a contaminated product.

<table>
<thead>
<tr>
<th>PersonID</th>
<th>Lname</th>
<th>Fname</th>
<th>Shipping_Address1</th>
<th>Probability</th>
</tr>
</thead>
<tbody>
<tr>
<td>108959616</td>
<td>Greenberg</td>
<td>Andrea</td>
<td>190 Maywood Drive</td>
<td>0.8</td>
</tr>
<tr>
<td>114543430</td>
<td>Lowery</td>
<td>Megan</td>
<td>10 Cyril Magnin St.</td>
<td>1</td>
</tr>
<tr>
<td>115753140</td>
<td>Tew</td>
<td>Quinn</td>
<td>301 pine st</td>
<td>1</td>
</tr>
</tbody>
</table>
query 2: product profitability

why?

Selling Price – (Cost of Ingredients + Cost of Label/Bottle + Cost of Employee Time)
query 2: product profitability

COST-BENEFIT

Generate list and graph profitability of each product over the span of a year to account for seasonal changes.
query 2: product profitability

Step 1: Find the cost of all single ingredient batches.

SELECT SIB.SI_BID AS SI_BID, ((SIB.Original_Weight*PRN.Cost) + E.Hourly_Wage*(EM.Time_Ended - EM.Time_Started)/60) AS Cost
FROM EMPLOYEE AS E, PRODUCT_RECORD_NONSELLABLE AS PRN, SINGLE_INGREDIENT_BATCH AS SIB, EMPLOYEE_MAKES_SIBATCH_WITH_EQUIP AS EM
WHERE PRN.PID = SIB.Ingredient_PID AND EM.BID = SIB.SI_BID AND EM.EID = E.EID;
query 2: product profitability

Step 2: Find the cost of all mixed ingredient batches.

```
SELECT MB.Mixed_BID AS Mix_BID, SUM((BCB.Liquid_Volume/SIB.Produced_Volume)*CSB.Cost) AS Cost
FROM MIXED_BATCH AS MB, MBATCH_CONTAINS_SIBATCH AS BCB, costSingleBatch AS CSB,
SINGLE_INGREDIENT_BATCH AS SIB
WHERE MB.Mixed_BID = BCB.Mixed_BID and
  BCB.SI_BID = CSB.SI_BID
GROUP BY MB.Mixed_BID;
```
Step 3: Find the profit of each product instance, grouped by product.

SELECT PRS.Pdate, PRS.PID, (PRS.Price-((B.Size/MBB.Final_Volume)*(MBC.Cost))) AS Profit
FROM PRODUCT_RECORD_SELLABLE AS PRS, BOTTLE AS B, MIXED_BATCH AS MBB, costMixBatch AS MBC, Prod_Includes_Prod AS PIP
WHERE PRS.PID = PIP.Product AND
  B.PID = PIP.Sub_Product AND
  PRS.Mixed_BID1 = MBB.Mixed_BID AND
  MBB.Mixed_BID = MBC.Mix_BID;
query 2: product profitability

Step 4: Total profit of each product.

SELECT PP.PDate, PP.PID, Sum(PP.Profit)
FROM ProductProfit AS PP
GROUP BY PP.PID, PP.Pdate;
query 3: sampling events

why?

WHAT MAKES IT SUCCESSFUL?

CUSTOMERS?

SAMPLES?

TYPE?
query 3: sampling events

Selection of information regarding each sampling and testing the correlation between these factors and the success of the event (customer yield).
query 3: sampling events

Step 1: Generates which sampling events potential customers attended and which products they tasted.

```sql
SELECT j.pid, sum(IIf(pc1.contactID = c.cid AND pc1.sampling_event = e1.eventID,1,0))
AS New_Customers, e1.eventID
FROM potential_customer AS pc1, sampling_event AS e1, customer AS c, juices_at_event AS j
WHERE (((pc1.sampling_event)=e1.eventID) And ((pc1.contactID) = c.cid) and ((e1.eventID)=j.eventID))
GROUP BY j.pid, e1.eventID;
```
query 3: sampling events

Step 2: Determines customer yield depending on type and product sampled.

```sql
SELECT j.pid, Q3.New_Customers, Count(pc2.contactID) AS Event_Attendees,
FROM potential_customer AS pc2, sampling_event AS e1, juices_at_event AS j, query3 AS Q3
WHERE pc2.sampling_event = e1.eventID AND e1.eventID=j.eventID
GROUP BY j.pid, e1.eventID, e1.type, Q3.New_Customers, e1.event_description;
```
query 3: sampling events

Multivariate regression in Excel with both product sampled and event type considered shows us that the product does not effect the yield, however…

![Regression Statistics Table]

<table>
<thead>
<tr>
<th>Regression Statistics</th>
</tr>
</thead>
<tbody>
<tr>
<td>Multiple R</td>
</tr>
<tr>
<td>R Square</td>
</tr>
<tr>
<td>Adjusted R Square</td>
</tr>
<tr>
<td>Standard Error</td>
</tr>
<tr>
<td>Observations</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>ANOVA</th>
</tr>
</thead>
<tbody>
<tr>
<td>df</td>
</tr>
<tr>
<td>Regression</td>
</tr>
<tr>
<td>Residual</td>
</tr>
<tr>
<td>Total</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Coefficients</th>
<th>Standard Error</th>
<th>t Stat</th>
<th>P-Value</th>
<th>Lower 95%</th>
<th>Upper 95%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>0.131944444</td>
<td>0.169653934</td>
<td>0.778</td>
<td>0.49374</td>
<td>-0.23168</td>
</tr>
<tr>
<td>1003</td>
<td>3.00964E-17</td>
<td>0.251467426</td>
<td>1.22E-16</td>
<td>1</td>
<td>0.539344</td>
</tr>
<tr>
<td>1006</td>
<td>4.65044E-17</td>
<td>0.251467426</td>
<td>1.66E-16</td>
<td>1</td>
<td>0.539344</td>
</tr>
<tr>
<td>1008</td>
<td>4.30555E-17</td>
<td>0.251467426</td>
<td>1.67E-16</td>
<td>1</td>
<td>0.539344</td>
</tr>
<tr>
<td>1009</td>
<td>1.04547E-16</td>
<td>0.214452142</td>
<td>5.15E-16</td>
<td>1</td>
<td>0.459954</td>
</tr>
<tr>
<td>1011</td>
<td>1.40765E-17</td>
<td>0.214452142</td>
<td>4.48E-16</td>
<td>1</td>
<td>0.459954</td>
</tr>
<tr>
<td>1013</td>
<td>1.04765E-17</td>
<td>0.214452142</td>
<td>3.99E-16</td>
<td>1</td>
<td>0.563333</td>
</tr>
<tr>
<td>1014</td>
<td>5.65112E-17</td>
<td>0.214452142</td>
<td>2.11E-16</td>
<td>1</td>
<td>0.563333</td>
</tr>
<tr>
<td>1016</td>
<td>0</td>
<td>0</td>
<td>66</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>2</td>
<td>0.263889309</td>
<td>0.214452142</td>
<td>1.23E-26</td>
<td>1</td>
<td>0.19607</td>
</tr>
<tr>
<td>3</td>
<td>0.065792222</td>
<td>0.251467426</td>
<td>0.262</td>
<td>1</td>
<td>0.796873</td>
</tr>
<tr>
<td>4</td>
<td>0.065792222</td>
<td>0.251467426</td>
<td>0.262</td>
<td>1</td>
<td>0.796873</td>
</tr>
<tr>
<td>5</td>
<td>0.263889309</td>
<td>0.214452142</td>
<td>1.23E-26</td>
<td>1</td>
<td>0.19607</td>
</tr>
</tbody>
</table>
query 3: sampling events

Another regression shows us that the **type of event** does impact the new customer yield.

Event types 2 (food events such as farmers’ markets) and 5 (health care related sampling events) produce the maximum yields (26.4% each)
query 4: custom cleanse

They have 3 cleanse options.

And one “Custom” option.

What are substitutes are customers choosing?

How can we use this information to create a forecast of demand for the next week?
query 4: custom cleanse

Find which juices are typically chosen in custom cleanses over those in standard cleanse in order to determine a list of juices are most popular and which

CUSTOMER ANALYSIS
query 4: custom cleanse

Step 1: Find which juices are typically chosen in custom cleanses over those in a standard cleanse in order to determine a list of juices most likely to be replaced and which juices they are being replaced with.

```sql
select Sum(IIf(CD.Juicel = 1003, CD.Quantity, 0))/sum(CD.Quantity) AS DRG,
       Sum(IIf(CD.Juicel = 1004, CD.Quantity, 0))/sum(CD.Quantity) AS DRGPlus,
       Sum(IIf(CD.Juicel = 1005, CD.Quantity, 0))/sum(CD.Quantity) AS EZG,
       Sum(IIf(CD.Juicel = 1006, CD.Quantity, 0))/sum(CD.Quantity) AS EZGPlus,
       sum(CD.Quantity) AS n
from Cleanse_Details AS CD
where exists (select * from [Order] 0
  where 0.OID = CD.OID and Date_Placed between #04/15/2013# AND #04/21/2013#);
```
Step 2: Forecast the next week's demand for the four juices using an exponential moving average, $\alpha = 0.5$, $t = 3$

Forecasted demand $= \sum_{i=0}^{t-1} \alpha(1 - \alpha)^i D_{t-i}$

```sql
SELECT (0.5*W3.DRG*W3.n + 0.25*W2.DRG*W2.n + 0.125*W1.DRG*W1.n)/
(0.5*W3.n + 0.25*W2.n + 0.125*W1.n) AS DRG,
(0.5*W3.DRGPlus*W3.n + 0.25*W2.DRGPlus*W2.n + 0.125*W1.DRGPlus*W1.n)/
(0.5*W3.n + 0.25*W2.n + 0.125*W1.n) AS DRGPlus,
(0.5*W3.EZG*W3.n + 0.25*W2.EZG*W2.n + 0.125*W1.EZG*W1.n)/
(0.5*W3.n + 0.25*W2.n + 0.125*W1.n) AS EZG,
(0.5*W3.EZGPlus*W3.n + 0.25*W2.EZGPlus*W2.n + 0.125*W1.EZGPlus*W1.n)/
(0.5*W3.n + 0.25*W2.n + 0.125*W1.n) AS EZGPlus,
(0.5*W3.n + 0.25*W2.n + 0.125*W1.n) AS n
FROM Week1Trend AS W1, Week2Trend AS W2, Week3Trend AS W3;
```
query 4: custom cleanse

Exponential Moving Average

[Graph showing exponential moving averages for different categories from 4/15-4/21 to 4/29-5/05, with a forecast for the future]
query 5: juice yield analysis

why?

BAGS

BLADES

DOES THE BAG OR BLADE MATTER?
query 5: juice yield analysis

Determine how much juice different blade (slice raw produce) and bag (strain juice) combinations generate.

OPTIMIZATION and HYPOTHESIS TESTING
query 5: juice yield analysis

Determine how much juice different blade (slice raw produce) and bag (strain juice) combinations generate.

SELECT s.BagID, s.BladeID, 
Avg(s.Produced_Volume/s.Original_Weight) AS [Yield(Mean)],
STDEV(s.Produced_Volume/s.Original_Weight) AS [Standard Dev],
COUNT(s.si_bid) AS n
FROM single_ingredient_batch AS s
WHERE (s.ingredient_PID=3001) And s.Date_Made Between #1/1/2012# And #12/30/2012#
GROUP BY s.BagID, s.BladeID
ORDER BY Avg(s.Produced_Volume/s.Original_Weight) DESC;
query 5: juice yield analysis

Determine how much juice different blade (slice raw produce) and bag (strain juice) combinations generate

<table>
<thead>
<tr>
<th>BagID</th>
<th>BladeID</th>
<th>Yield</th>
</tr>
</thead>
<tbody>
<tr>
<td>9501</td>
<td>9601</td>
<td>0.1136363664</td>
</tr>
<tr>
<td>9501</td>
<td>9601</td>
<td>0.141025641</td>
</tr>
<tr>
<td>9501</td>
<td>9602</td>
<td>0.126760563</td>
</tr>
<tr>
<td>9501</td>
<td>9602</td>
<td>0.147058824</td>
</tr>
<tr>
<td>9502</td>
<td>9601</td>
<td>0.24</td>
</tr>
<tr>
<td>9502</td>
<td>9601</td>
<td>0.37037037</td>
</tr>
<tr>
<td>9502</td>
<td>9602</td>
<td>0.178571429</td>
</tr>
<tr>
<td>9502</td>
<td>9602</td>
<td>0.215384615</td>
</tr>
</tbody>
</table>
query 5: juice yield analysis

Use MS Excel to generate two factor ANOVA with Replication

<table>
<thead>
<tr>
<th>Source of Variation</th>
<th>SS</th>
<th>df</th>
<th>MS</th>
<th>F</th>
<th>P-value</th>
<th>F crit</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bag</td>
<td>0.028304</td>
<td>1</td>
<td>0.028304</td>
<td>11.60348</td>
<td>0.027117</td>
<td>7.708647</td>
</tr>
<tr>
<td>Blade</td>
<td>0.004864</td>
<td>1</td>
<td>0.004864</td>
<td>1.933965</td>
<td>0.23078</td>
<td>7.708647</td>
</tr>
<tr>
<td>Interaction</td>
<td>0.005937</td>
<td>1</td>
<td>0.005937</td>
<td>2.84383</td>
<td>0.167003</td>
<td>7.708647</td>
</tr>
<tr>
<td>Within</td>
<td>0.009757</td>
<td>4</td>
<td>0.002439</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>0.049861</td>
<td>7</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
query 5: juice yield analysis

Determine that bag choice is significant, the choice of blade is not.

Conclusion: It does not matter which blade we use, it is important to use bag 9502 when pressing ingredient 3001 (carrots). This will increase yield, decrease waste, and minimize production time.
Normalization analysis

On tables Cleanse and Customer

<table>
<thead>
<tr>
<th>Partial Dependence</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Cleanse(PID\textsuperscript{12}, Cleanse_Name, Number_of_Days, Price)</td>
</tr>
<tr>
<td>1. CleansePrice(PID\textsuperscript{12}, Number_of_Days, Price)</td>
</tr>
<tr>
<td>2. CleanseName(PID\textsuperscript{12}, Cleanse_Name)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Transitive Dependence</th>
</tr>
</thead>
<tbody>
<tr>
<td>2. Customer(CID\textsuperscript{1}, Customer_Type, Group_Name, Bdate, Gender, Company, County, Tax_Rate, Notes)</td>
</tr>
<tr>
<td>1. Customer(CID\textsuperscript{1}, Customer_Type, Group_Name, Bdate, Gender, Company, County, Notes)</td>
</tr>
<tr>
<td>2. County(County, Tax_Rate)</td>
</tr>
</tbody>
</table>
Thank you!