	Lehrveranstaltung	Grundlagen von Datenbanken			WS 2019/20
	Aufgabenzettel	03			
	STiNE-Gruppe 01	Dahmen, Zdravkovic, Ostendorf, Ludewigs			
	Ausgabe	Mi. 15.11.2019	Abgabe	Do. 29.11.2019	

1 Präsenzaufgabe

2 Präsenzaufgabe

3 Übungsaufgabe: Abbildung eines ER-Diagramms auf das relationale Datenmodell

Person (Pnr, Vorname, Nachname, Geb.)

BibliotekarIn(Person \Rightarrow Person.Pnr, Lieblingsgenre)

ProfessorIn (Person \Rightarrow Person.Pnr, Fakultaet, Spezialgebiet)

StudentIn (Person \Rightarrow Person.Pnr, Studienfach, #Fachsemester)

Bibliothek (Adresse, Name, #Mitarbeiter, Leiter \Rightarrow BibliotekarIn.Person)

Buch (ISBN, Titel, Preis, #Seiten, Autor \Rightarrow ProfessorIn.Person)

Beinhaltet (Buch \Rightarrow Buch.ISBN, Bibliothek \Rightarrow Bibliothek.Adresse, Bestand)

Empfiehl (StudentIn \Rightarrow StudentIn.Person, ProfessorIn \Rightarrow ProfessorIn.Person, Buch \Rightarrow Buch.ISBN) LeihtAus (StudentIn \Rightarrow StudentIn.Person, Buch \Rightarrow Buch.ISBN, Ausleihdatum, Rueckgabedatum)

4 Übungsaufgabe: Relationenalgebra

a)

π Name, Fachgebiet (σ HID = Handwerker (σ Firmensitz = Immobilie (Handwerksfirma \bowtie Auftrag)))

b)


$(\pi$ EID, Vorname, Nachname Eigentuerer) - $(\pi$ Vermieter, Vorname, Nachname (Eigentuerer \bowtie Auftrag))

c)

π IID, Adresse (σ Baujahr \geq 1980 (σ Bauherr \neq Eigentuerer Immobilie))

d)

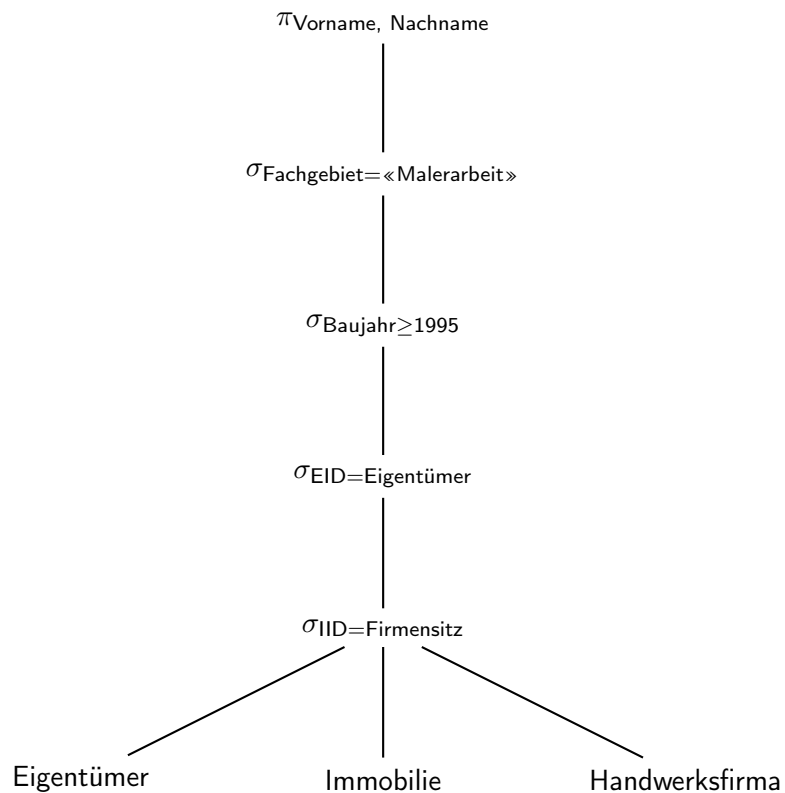
Wir suchen die Adresse und das Baujahr von Immobilien und den Nachnamen von dessen Eigentuern, welche in einem Auftrag Kosten von 10.000 oder weniger hinterlegt bekommen haben.


	Lehrveranstaltung	Grundlagen von Datenbanken			WS 2019/20
	Aufgabenzettel	03			
	STiNE-Gruppe 01	Dahmen, Zdravkovic, Ostendorf, Ludewigs			
	Ausgabe	Mi. 15.11.2019	Abgabe	Do. 29.11.2019	

5 Übungsaufgabe: Algebraische Optimierung

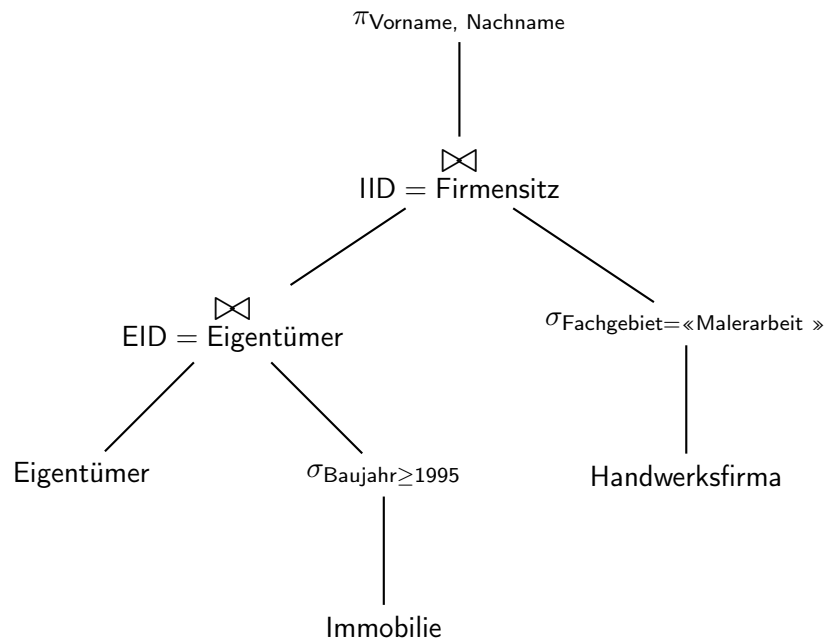
Gesucht sind Vor- und Nachname all derer Eigentümer, die eine Immobilie besitzen, die sowohl von einer auf Malerei spezialisierten Handwerksfirma als Firmensitz verwendet wird, als auch frühestens im Jahre 17975 erbaut wurde.

a)



	Lehrveranstaltung	Grundlagen von Datenbanken			WS 2019/20
	Aufgabenzettel	03			
	STiNE-Gruppe 01	Dahmen, Zdravkovic, Ostendorf, Ludewigs			
	Ausgabe	Mi. 15.11.2019	Abgabe	Do. 29.11.2019	

b)



Optimierungsgrad & Heuristiken

5.0.1 Führe Selektionen so früh wie möglich aus

Die Anfrage in b) versucht nach jedem JOIN direkt eine Selektion auszuführen um die Anzahl an Zwischenergebnissen direkt zu reduzieren. In a) werden hingegen zunächst JOINS ausgeführt und dabei unnötige Abfragen erstellt, die vermieden werden könnten.


5.0.2 Führe Projektionen so früh wie möglich aus

Sowohl a) als auch b) führen die einzige Projektion so früh wie möglich aus. Leider kann in diesem Fall die Projektion erst nach allen JOINS ausgeführt werden.

Dies ist wichtig um die Zwischenergebnisse so klein wie möglich zu halten.

5.0.3 Verknüpfe Folgen von unnären Operatoren

Die Anfrage in a) führt die JOINS zu einem Zeitpunkt aus und kann so mehrere Ergebnisse zusammenfassen. Da aber alle Ergebnisse voll erweitert werden ist dies kein Vorteil.

	Lehrveranstaltung	Grundlagen von Datenbanken			WS 2019/20
	Aufgabenzettel	03			
	STiNE-Gruppe 01	Dahmen, Zdravkovic, Ostendorf, Ludewigs			
	Ausgabe	Mi. 15.11.2019	Abgabe	Do. 29.11.2019	

5.0.4 Fass einfache Selektionen auf einer Relation zusammen

In a) werden alle Selektionen nacheinander ausgeführt. Dadurch sind unnötig viele Schritte notwendig obwohl die Information zum Wegstreichen von mehreren Reihen zur Verfügung steht.

5.0.5 Verknüpfe bestimmte Selektionen mit einem vorausgehenden Kartesischen Produkt zu einem Verbund

In a) wird zwar ein kartesisches Produkt gebildet, die Selektionen werden aber nicht im Verbund durchgeführt. b) hat gar keine Produkte

5.0.6 Berechne gemeinsame Teilbäume nur einmal

Beide Anfragen verwenden keine Teilbäume doppelt

5.0.7 Fazit

Die Anfrage in a) sammelt als erstes so viele Daten wie möglich und führt die Selektionen einzeln aus. Dieses Vorgehen belegt viel Speicher und führt zu langsamen Selektionen. Der einzige Vorteil ist die einfache Generierung und Parallelisierbarkeit der Anfrage. In b) wird die Ergebnismenge erst gekürzt und dann erweitert. Dies führt zu kurzen Selektionsvorgängen und kleinen Datenmengen im Speicher.

Da einige Anfragen aber parallel ausgeführt werden, ist die gesamte Anfrage nur schlecht parallelisierbar.

Insgesamt hat b) einen höheren Optimierungsgrad