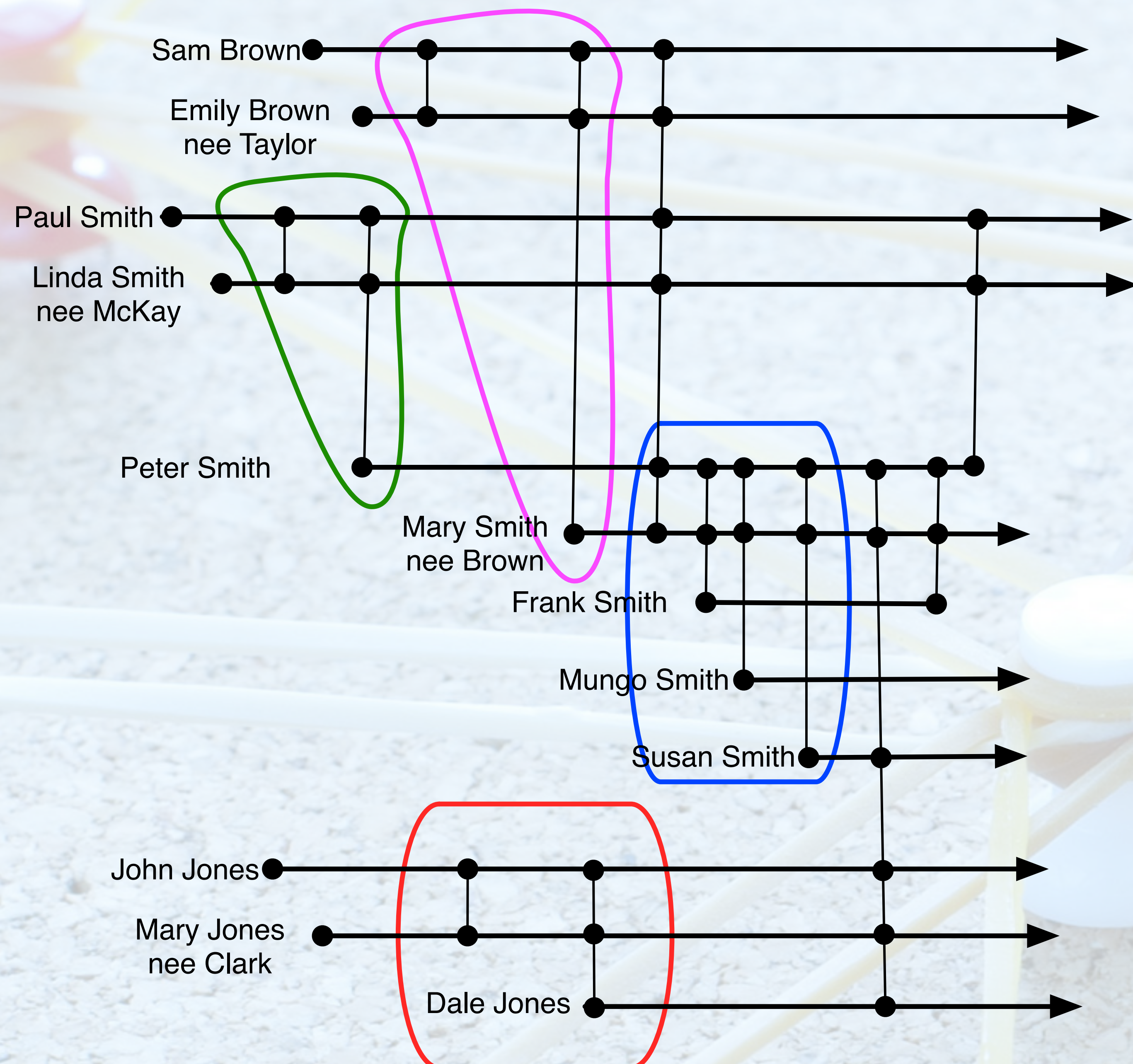


Linkage of Births, Marriages and Deaths

Özgür Akgün, Peter Christen, Tom Dalton, Alan Dearle, Chris Dibben, Eilidh Garrett, Graham Kirby

The Problem

The aim of our research is to take a large set of *Birth, Marriage and Death records* and automatically link them to create a **genealogical pedigree**. The challenges of this task are due to the lack of unique identifiers and the level of data loss and corruption. We make use of **Probabilistic Data Linkage** approaches to identify **conjugal family units**. This differs from most linkage techniques which focus primarily on individuals.



The Approach

The first technique that is employed is to **identify familial groups**. This attempts to gather siblings' birth records and their parents' marriage record into bundles with the aim of partitioning the certificates into birth-marriage familial groups.

Bundling relies on the observation that Scottish Birth and Marriage records record the names of parents and children along with the date and place of marriage of the parents.

Familial groups are reorganised into sets of potential families indexed using the names of the parents.

We **match** each death record to potential families. Matches are found by comparing the names and date of births on the death record with records in the sets of familial groups. If a match is found a new death-marriage familial group is formed.

We **merge birth-marriage** and **death-marriage** familial groups based on the shared marriage record and the spacing of children.

At this stage the groups still do not contain any child marriages and are thus independent familial groups rather than linked pedigrees.

To make **inter-generational links** we use marriage records in a familial group to identify which set of potential families to find the parents' birth records. Based on comparisons of the marriage record to the set of potential families inter-generational links can be made.

Gold-standard Evaluation

- **Evaluating the success of linkage algorithms is hard.**
 - For real-world data, the true linkage is unknown.
 - If the true links were known, we would not be linking it.
- To evaluate
 - We use manually linked data sets (from Skye and Kilmarnock)
- Shortcomings
 - Manual links may be wrong
 - Small number of data sets, overfitting
 - Limited size of data sets, hard to test scaling
- For robust evaluation, we create **synthetic longitudinal genealogical** data sets
 - True links are known
 - Large number of data sets with varying characteristics, avoiding overfitting

The Input Data

Birth, Marriage, Death vital event record from 1856-1973 for Scotland

Errors typically include:

- Missing data, typos
- Spelling variations
- Nicknames instead of full names
- Transcription errors
- Translation / transliteration errors

- *Birth certificates* record:
 - the names of the parents,
 - the date and place of their marriage.
- *Marriage certificates* record:
 - the name of the groom and the bride,
 - the date and place of their marriage.

Outputs

- The desired output of linkage depends on the intended use of the linked data.
- We record **provenance** information throughout our linkage process for each potential link.
- A final pedigree can be projected
 - Using the recorded provenance
 - By selecting links that meet a desired confidence threshold
 - Satisfying data restrictions, or other domain-specific rules
- We treat linkage as a continual process.
 - Recording provenance of potential links allows easier addition of new data.

Page 2]

The undermentioned Houses are situate within the Boundaries of the

Civil Parish of

Bracadale

Municipal Borough of

Quasi-Sovere Parish of

School Board District of

Bracadale

Parliamentary Borough of

Parliamentary Division of

Royal Borough of

Police-Borough of

Borough Ward of

Town of

Village or Hamlet of

Island of

Bracadale

No. of Houses.	ROAD, STREET, &c., and No. or NAME of HOUSE.	HOUSE No. (as shown on map)	NAME and Surname of each Person.	RELATION to Head of Family.	CONNECTION as to Marriage.	AGE (last birthday)	PROFESSION or OCCUPATION.	Employer.	Employed.	WHERE BORN.	Gender, or G.D.R.	Number of Children, or issue, living, or dead.	Notes with reference to House.
4	1	1	James Brown	Head	Married	34	Labourer			Bracadale	M	2	
5	2	1	John Brown	Wife	Married	32	Labourer			Bracadale	F	1	
	3	1	William Brown	Son	Single	10	Schoolboy			Bracadale	M	0	
	4	1	Elizabeth Brown	Daughter	Single	8	Schoolgirl			Bracadale	F	0	
	5	1	James Brown	Son	Single	6	Schoolboy			Bracadale	M	0	
	6	1	William Brown	Son	Single	4	Schoolboy			Bracadale	M	0	
	7	1	Elizabeth Brown	Daughter	Single	2	Schoolgirl			Bracadale	F	0	
6	1	1	James Brown	Head	Married	34	Labourer			Bracadale	M	2	
7	2	1	John Brown	Wife	Married	32	Labourer			Bracadale	F	1	
8	3	1	William Brown	Son	Single	10	Schoolboy			Bracadale	M	0	
	4	1	Elizabeth Brown	Daughter	Single	8	Schoolgirl			Bracadale	F	0	
	5	1	James Brown	Son	Single	6	Schoolboy			Bracadale	M	0	
	6	1	William Brown	Son	Single	4	Schoolboy			Bracadale	M	0	
	7	1	Elizabeth Brown	Daughter	Single	2	Schoolgirl			Bracadale	F	0	
	8	1	James Brown	Son	Single	0	Schoolboy			Bracadale	M	0	
	9	1	William Brown	Son	Single	0	Schoolboy			Bracadale	M	0	
	10	1	Elizabeth Brown	Daughter	Single	0	Schoolgirl			Bracadale	F	0	
	11	1	James Brown	Son	Single	0	Schoolboy			Bracadale	M	0	
	12	1	William Brown	Son	Single	0	Schoolboy			Bracadale	M	0	
	13	1	Elizabeth Brown	Daughter	Single	0	Schoolgirl			Bracadale	F	0	
	14	1	James Brown	Son	Single	0	Schoolboy			Bracadale	M	0	
	15	1	William Brown	Son	Single	0	Schoolboy			Bracadale	M	0	
	16	1	Elizabeth Brown	Daughter	Single	0	Schoolgirl			Bracadale	F	0	
	17	1	James Brown	Son	Single	0	Schoolboy			Bracadale	M	0	
	18	1	William Brown	Son	Single	0	Schoolboy			Bracadale	M	0	
	19	1	Elizabeth Brown	Daughter	Single	0	Schoolgirl			Bracadale	F	0	
	20	1	James Brown	Son	Single	0	Schoolboy			Bracadale	M	0	
	21	1	William Brown	Son	Single	0	Schoolboy			Bracadale	M	0	
	22	1	Elizabeth Brown	Daughter	Single	0	Schoolgirl			Bracadale	F	0	
	23	1	James Brown	Son	Single	0	Schoolboy			Bracadale	M	0	
	24	1	William Brown	Son	Single	0	Schoolboy			Bracadale	M	0	
	25	1	Elizabeth Brown	Daughter	Single	0	Schoolgirl			Bracadale	F	0	
	26	1	James Brown	Son	Single	0	Schoolboy			Bracadale	M	0	
	27	1	William Brown	Son	Single	0	Schoolboy			Bracadale	M	0	
	28	1	Elizabeth Brown	Daughter	Single	0	Schoolgirl			Bracadale	F	0	
	29	1	James Brown	Son	Single	0	Schoolboy			Bracadale	M	0	
	30	1	William Brown	Son	Single	0	Schoolboy			Bracadale	M	0	
	31	1	Elizabeth Brown	Daughter	Single	0	Schoolgirl			Bracadale	F	0	
	32	1	James Brown	Son	Single	0	Schoolboy			Bracadale	M	0	
	33	1	William Brown	Son	Single	0	Schoolboy			Bracadale	M	0	
	34	1	Elizabeth Brown	Daughter	Single	0	Schoolgirl			Bracadale	F	0	
	35	1	James Brown	Son	Single	0	Schoolboy			Bracadale	M	0	
	36	1	William Brown	Son	Single	0	Schoolboy			Bracadale	M	0	
	37	1	Elizabeth Brown	Daughter	Single	0	Schoolgirl			Bracadale	F	0	
	38	1	James Brown	Son	Single	0	Schoolboy			Bracadale	M	0	
	39	1	William Brown	Son	Single	0	Schoolboy			Bracadale	M	0	
	40	1	Elizabeth Brown	Daughter	Single	0	Schoolgirl			Bracadale	F	0	
	41	1	James Brown	Son	Single	0	Schoolboy			Bracadale	M	0	
	42	1	William Brown	Son	Single	0	Schoolboy			Bracadale	M	0	
	43	1	Elizabeth Brown	Daughter	Single	0	Schoolgirl			Bracadale	F	0	
	44	1	James Brown	Son	Single	0	Schoolboy			Bracadale	M	0	
	45	1	William Brown	Son	Single	0	Schoolboy			Bracadale	M	0	
	46	1	Elizabeth Brown	Daughter	Single	0	Schoolgirl			Bracadale	F	0	
	47	1	James Brown	Son	Single	0	Schoolboy			Bracadale	M	0	
	48	1	William Brown	Son	Single	0	Schoolboy			Bracadale	M	0	
	49	1	Elizabeth Brown	Daughter	Single	0	Schoolgirl			Bracadale	F	0	
	50	1	James Brown	Son	Single	0	Schoolboy			Bracadale	M	0	
	51	1	William Brown	Son	Single	0	Schoolboy			Bracadale	M	0	
	52	1	Elizabeth Brown	Daughter	Single	0	Schoolgirl			Bracadale	F	0	
	53	1	James Brown	Son	Single	0	Schoolboy			Bracadale	M	0	
	54	1	William Brown	Son	Single	0	Schoolboy			Bracadale	M	0	
	55	1	Elizabeth Brown	Daughter	Single	0	Schoolgirl			Bracadale	F	0	
	56	1	James Brown	Son	Single	0	Schoolboy			Bracadale	M	0	
	57	1	William Brown	Son	Single	0	Schoolboy			Bracadale	M	0	
	58	1	Elizabeth Brown	Daughter	Single	0	Schoolgirl			Bracadale	F	0	
	59	1	James Brown	Son	Single	0	Schoolboy			Bracadale	M	0	
	60	1	William Brown	Son	Single	0	Schoolboy			Bracadale	M	0	
	61	1	Elizabeth Brown	Daughter	Single	0	Schoolgirl			Bracadale	F	0	
	62	1	James Brown	Son	Single	0	Schoolboy			Bracadale	M	0	
	63	1	William Brown	Son	Single	0	Schoolboy			Bracadale	M	0	
	64	1	Elizabeth Brown	Daughter	Single	0	Schoolgirl			Bracadale	F	0	
	65	1	James Brown	Son	Single	0	Schoolboy			Bracadale	M	0	
	66	1	William Brown	Son	Single	0	Schoolboy			Bracadale	M	0	
	67	1	Elizabeth Brown	Daughter	Single	0	Schoolgirl			Bracadale	F	0	
	68	1	James Brown	Son	Single	0	Schoolboy			Bracadale	M	0	
	69	1	William Brown	Son	Single	0	Schoolboy			Bracadale	M	0	
	70	1	Elizabeth Brown	Daughter	Single	0	Schoolgirl			Bracadale	F	0	
	71	1	James Brown	Son	Single	0	Schoolboy			Bracadale	M	0	
	72	1	William Brown	Son	Single	0	Schoolboy			Bracadale	M	0	
	73	1	Elizabeth Brown	Daughter	Single	0	Schoolgirl			Bracadale	F	0	
	74	1	James Brown	Son	Single	0	Schoolboy			Bracadale	M	0	
	75	1	William Brown	Son	Single	0	Schoolboy			Bracadale	M	0	
	76	1	Elizabeth Brown	Daughter	Single	0	Schoolgirl			Bracadale	F	0	
	77	1	James Brown	Son	Single	0	Schoolboy			Bracadale	M	0	
	78	1	William Brown	Son	Single	0	Schoolboy			Bracadale	M	0	
	79	1	Elizabeth Brown	Daughter	Single	0	Schoolgirl			Bracadale	F	0	
	80	1	James Brown	Son	Single	0	Schoolboy			Bracadale	M	0	
	81	1	William Brown	Son	Single	0	Schoolboy			Bracadale	M	0	
	82	1	Elizabeth Brown	Daughter	Single	0	Schoolgirl			Bracadale	F	0	
	83	1	James Brown	Son	Single	0	Schoolboy			Bracadale	M	0	
	84	1	William Brown	Son	Single	0	Schoolboy			Bracadale	M	0	
	85	1	Elizabeth Brown	Daughter	Single	0	Schoolgirl			Bracadale	F	0	
	86	1	James Brown	Son	Single	0	Schoolboy			Bracadale	M	0	
	87	1	William Brown	Son	Single	0	Schoolboy			Bracadale	M	0	
	88	1	Elizabeth Brown	Daughter	Single	0	Schoolgirl			Bracadale	F	0	
	89	1	James Brown	Son	Single	0	Schoolboy			Bracadale	M	0	
	90	1	William Brown	Son	Single	0	Schoolboy			Bracadale	M	0	
	91	1	Elizabeth Brown	Daughter	Single	0	Schoolgirl			Bracadale	F	0	
	92	1	James Brown	Son	Single	0	Schoolboy			Bracadale	M	0	
	93	1	William Brown	Son	Single	0	Schoolboy			Bracadale	M	0	
	94	1	Elizabeth Brown	Daughter	Single	0	Schoolgirl			Bracadale	F	0	
	95	1	James Brown	Son	Single	0	Schoolboy			Bracadale	M	0	
	96	1	William Brown	Son	Single	0	Schoolboy			Bracadale	M	0	
	97	1	Elizabeth Brown	Daughter	Single	0	Schoolgirl			Bracadale	F	0	
	98	1	James Brown	Son	Single	0	Schoolboy			Bracadale	M	0	
	99	1	William Brown	Son	Single	0	Schoolboy			Bracadale	M	0	
	100	1	Elizabeth Brown	Daughter	Single	0	Schoolgirl			Bracadale	F	0	

Scor.—Sheet A.]

Name—District—

Total of
Windows
& Rooms.

27